DOCUMENT RESUME

ED 133 066	PS 009 006

AUTHOR	Stodolsky, Susan S.
TITLE	Ancona Montessori Research Project for Culturally Disadvantaged Children. Final Report.
INSTITUTION	Chicago Univ., Ill.
SPONS AGENCY	Office of Economic Opportunity, Washington, D.C.
PUB DATE	31 Aug 70
GRANT	OEO-CG-8469-A/2
NOTE	125p.
EDRS PRICE	MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS	Academic Achievement; Cognitive Development; *Culturally Disadvantaged; Curriculum Evaluation; Disadvantaged Youth; *Early Childhood Education; *Educational Objectives; Elementary School Students; Evaluation Methods; Intelligence Tests; Intervention; Longitudinal Studies; Middle Class; *Nursery Schools; *Parent Participation; Preschool Children; *Program Evaluation; Psychomotor Skills; Rating Scales; Social Development; Social Relations
IDENTIFIERS	*Montessori Method

ABSTRACT

This is the final report of the Ancona Montessori Research Project for Culturally Disadvantaged Children begun in 1965 to investigate the effects of a modified Montessori program for disadvantaged children in the preschool and early elementary years. This report deals with the academic year 1969-1970, in which 29 disadvantaged children and a comparable group of 29 middle class children are the central focus of study. In addition, there is a followup on the school careers of disadvantaged children who attended Ancona at one time. A number of hypotheses about the potential effects of the project on the children's cognitive, social development are studied. Part I of the report deals with findings relative to the nursery school children, and includes a discussion of data from three measures of intellectual development (Stanford Binet, WPPSI and Merrill-Palmer) and from tester and teacher ratings of school-related behaviors and attitudes and social interaction. Part II details findings on the elementary school children and followup data on children who attended Ancona in previous years but are now elementary school students in other schools. In addition, data regarding children whose families have had long term involvement in the school is discussed. The appendix includes Ancona school Head Start program ratings of behavior during individual intelligence testing. (MS)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every

11 76

ps

ED133066

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

FINAL REPORT: OEO GRANT CG-8469 A/2

Submitted by: Susan S. Stodolsky, Principal Investigator
Alfred Karlson, Research Director

Title of Project: Ancona Montessori Research Project for Culturally Disadvantaged Children

Reporting Period: September 1, 1969 to August 31, 1970

PS 009006

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS i

INTRODUCTION 1

PART I: THE STUDY OF NURSERY SCHOOL CHILDREN 5

RESULTS AND DISCUSSION 13
- Stanford-Binet Data 13
- WPPSI Data 20
- Merrill-Palmer Data 28
- Test and Teacher Ratings 30
- Birch Procedure 40
- Social Interaction 53

PART II: THE STUDY OF ELEMENTARY SCHOOL CHILDREN AND SIBLINGS WITHIN FAMILIES WITH CONTINUED ENROLLMENT 68

RESULTS AND DISCUSSION 71
- Public School Children 71
- Ancona Elementary Children 75
  - Metropolitan Data 75
  - WISC Data 77
  - Teacher Ratings 80
- Diffusion Effects 83

OVERVIEW 89

APPENDIX A: Ancona School Headstart Program Ratings of Behavior During Individual Intelligence Testing

APPENDIX B: Report on the Social Work Program

ACKNOWLEDGEMENTS

The execution of the research reported here was facilitated by the cooperation of many persons at the Ancona School and The University of Chicago. The teachers and children at the school were very helpful in making themselves available to us and providing assistance and facilities when necessary. The parents of the Headstart children were also most cooperative in complying with requests for permission to follow and test children, even when they were no longer in the Ancona School. Lila Gordon, the Director of the School, helped us in many ways including administering tests, and sharing her insights about children and their parents with us. Gwen Washington assisted us in many ways including helping us keep track of children, each other, and finances.

The bulk of the field work which went into this study was supervised by Alfred Karlson who also did much testing and observation in the School. Nancy Miller took almost total responsibility for the social interaction study and has written the section of this report dealing with that area, she also administered tests. Ellen Potter assisted in both the testing program and the classroom observations. Christine Halfar participated in the observational part of the study. The work of this staff was much appreciated, not only for its general excellence and punctuality, but for the good spirit maintained by all involved. We are all indebted to Judy Jensen, who was available to us throughout the course of this work to remind us of things done in the past and help us think through our current procedures, and because previous research materials were all easily located and understood due to her organizational abilities.

Sue Gross and Spencer Swinton assisted us in getting the data through the computer with speed and good will. The work of our coders, Bob Jaffe, Jon Jaffe, Barbara Golter, and Martha Garber was also much appreciated. Finally, Kathy Cotton was able through her skill to turn our many scribblings into a final typed manuscript.

INTRODUCTION

Since 1965 the Ancona Montessori School has had OEO support to investigate the effects of a modified Montessori program for disadvantaged children in the preschool and early elementary years. This report deals with the academic year 1969-70. During this year thirty-nine disadvantaged children and a comparison group of twenty-nine middle-class children have been the central focus of study as they are currently attending the school. In addition, we have continued to follow the school careers of disadvantaged children who attended Ancona at one time or another since 1965 and are now attending school elsewhere.

Twenty-nine of the disadvantaged children who attended Ancona this year are in the age range from three to six years and participated in the nursery classes at the school. Ten children in the current sample were disadvantaged children who had completed the Ancona nursery program previously and were attending the elementary school classes. For ease of presentation, this report will be divided so that the first part will deal with the findings relative to the nursery school children. Part II will detail findings on the elementary school children and follow-up data on children who attended Ancona in previous years but are now elementary school children in other institutions. In addition, data regarding children whose families have had long-term involvement in the school will be discussed in Part II.

The Ancona Montessori School is a parent-governed nursery and primary school. The school serves a population of children who are predominantly of middle-class background but the racial composition of the school is quite balanced including both Negro and white children. The disadvantaged children attending the school are all Negro and most come from the neighborhood immediately adjacent to the school. It has always been the policy of the school to place the disadvantaged children in its regular classrooms with a small number of these children in each classroom. In this fashion the diversity of most classrooms has been enhanced.

The overall objectives of the program as originally stated in our OEO proposal are focused in four areas:

(1) Enhancement of the children's intellectual growth through exposure to the structure, materials and methods of a modified Montessori classroom. The modified Montessori classrooms provide (a) opportunity for the development of attention and independent task interest through their stress upon individual teaching and learning with self-chosen, self-correcting materials; (b) opportunity for the development of abstract concepts through the Montessori sensory materials which provide training in sensory discrimination, matching and seriation, and through supplementary materials promoting the classification of real objects; and (c) opportunity for expressive and dramatic play.

(2) Providing continuity of educational intervention by early entrance into the school, and by enabling the children to remain in the school through the primary grades.

(3) Providing a school setting integrated by race and social class, in order to (a) expose the disadvantaged children to an atmosphere of greater task orientation and to the use of standard English speech; (b) provide both social groups with the opportunity for contact through common endeavors; and (c) provide the

opportunity for interaction among the parents of the middle-class and disadvantaged children through the extensive parent participation which is part of the Ancona School's program.

(4) Continued involvement with the families of the disadvantaged children, through (a) recruitment of younger siblings into the program; (b) a social work program aimed especially at the promotion of self-help through more effective problem-solving in the families; and (c) provision of medical services.

General Statement of Hypotheses To Be Investigated

The hypotheses studied in this research project deal generally with the effects of attendance at the Montessori School on the intellectual development of disadvantaged children, the effects on a cluster of school-facilitating behaviors such as attention and task orientation, and the effects on social interaction patterns. In addition, particular interest has been paid to the progress of children from families who have participated in the school's program over a number of years.

More specifically, the hypotheses currently under study are as follows:

Hypothesis 1: A Montessori program will increase the basic cognitive and behavioral skills of disadvantaged children required for educational achievement. This effect will be greatest if started early (age three) and continued over an extensive period of time.

Hypothesis 1a: The disadvantaged children will show increased cognitive development. We especially anticipate improvement in number concepts, visual discrimination, classificatory skills and psycho-motor skills which are emphasized in the curriculum.

Hypothesis 1b: The disadvantaged children will show increased attentiveness to task demands.

Hypothesis 1c: Disadvantaged children who continue in the Montessori elementary school program will show higher school achievement than those who go to public school.

Hypothesis 2: Interaction between middle-class and disadvantaged children will increase as a function of experience in the program.

Hypothesis 2a: There will be more cross-group social acts in children with longer tenure at the school.

Hypothesis 2b: There will be more cross-group friendships among children with longer tenure at the school.

Hypothesis 3: Continuing involvement of the same families in the program will result in "diffusion effects" to the intellectual attainments of older and younger siblings.

Hypothesis 3a: In examining children who attended Ancona at one time, older siblings will show school attainment that is better than non-siblings in public schools.

Hypothesis 3b: Younger siblings entering the program will show higher initial cognitive and behavioral skills.

As hypothesis 1c is in regard to the elementary school children, it will be discussed in Part II of this report. To the extent that other hypotheses also deal with older children or Ancona graduates, discussion of them will also be deferred. In addition, Hypothesis 3 will be discussed in Part II.

The research rationale which led to the formulation of these hypotheses for study has been detailed in our original reports. Rather than repeat the rationale here, we will incorporate discussion of the relevant literature as we present the actual findings and discuss them.

PART I

THE STUDY OF THE NURSERY SCHOOL CHILDREN

The overall design of the study was a classic pre- and post-test paradigm. The disadvantaged children attending the nursery classes and a comparison group of middle-class children in the same classrooms were tested early in the academic year and close to the end of the year. As can be seen from the hypotheses, the primary interest was in ascertaining if change occurred on a number of characteristics of the children throughout the course of the academic year and as a result of the preschool experience. Since this is a long-term study, data from previous years on some of the children are available and will be used to speak to questions of long-term change.

The Sample

Twenty-nine disadvantaged children participated in the classes of the Ancona preschool. These children were all Negro children coming from poor families who live in the neighborhood adjacent to the school. In general, the families of the children meet OEO guidelines for participation in Headstart programs; thus the families have poverty line incomes and many are mother-only households.

Thirteen of the disadvantaged children were in attendance at Ancona for the first time this year. A special effort was made to recruit children from families who had children in attendance at Ancona in previous years. Seven of the thirteen new children were younger siblings of children who were currently attending Ancona or who had participated in the Headstart program in previous years. The average age of the thirteen first-year children was thirty-nine months as of October, 1969. There were seven girls and six boys in this group.

The sixteen other disadvantaged children attending nursery classes had attended Ancona in previous years. Fourteen children were in their second year of attendance, and two were third year. This group had an average age of fifty-five months and was composed of eight boys and eight girls.

To the extent that it was possible, every disadvantaged child was pair-matched with a middle-class child in his own classroom. The criteria for matching were sex, age, and prior attendance history. The middle-class comparison sample was composed of seventeen first-year students and twelve students in their second to third year in the school. The average age for the first-year group was forty-five months; for the second- to third-year group it was fifty-six months. There were sixteen girls and thirteen boys in the middle-class comparison group.

The comparison group was not as closely matched this year as has been possible in previous years. In particular the first year Headstart children are typically somewhat younger than the available middle-class comparison children.

Other difficulties have arisen in regard to the Headstart sample this year. Massive urban renewal is currently in progress in the neighborhood where the families live and the school is located. Consequently, a number of families have left the area. Whenever possible, the school made available busing arrangements to families. Nevertheless, in the first months of school, four Headstart children did stop attending the school. Replacement children were enrolled two to three months late, and middle-class comparison children for these replacement children had to be selected from children who had been in school since the start of the year.

In addition, one child had very erratic attendance and left the school near the end of the year. The family's whereabouts could not be determined so the child and her contrast child were dropped from the sample group.

Instruments and Procedures

Most data gathering procedures were administered twice to all children in the disadvantaged and middle-class comparison groups. Testing was carried out during the beginning of the school year and close to the end of the school year. Test administration was done by three trained testers. The testers, two women and one man, are all advanced graduate students in educational psychology. The testers are all Caucasian. All three testers administered the Stanford-Binet and WPPSI tests. One tester did all Merrill-Palmer and WISC tests. An additional trained observer was used in collecting the social interaction observations.

The procedures used can be classified as to the general type of variable being measured. One set of variables deals with the intellectual or cognitive development of the children. Another set of procedures is measuring certain school related behaviors thought to be facilitative of school success. Another set assesses social interaction variables.

Measures of intellectual development

(1) The Stanford-Binet Intelligence Test, Form L-M. On the average, children were administered the Stanford-Binet after four weeks of school attendance for the first testing and after thirty-one weeks for the second testing. Thus there was approximately a twenty-seven-week interval between first and second testings.

The Stanford-Binet was administered according to standard procedures by testers who were trained to the usual standards. In general, the test was given in one session of about one hour, but testing was terminated if the child was tired or uncooperative and was continued at another session.

(2) Wechsler Preschool and Primary Scale of Intelligence (WPPSI). All sample children four years of age and older received selected scales from the WPPSI. In addition, a few children under four who seemed able to respond were also tested. The WPPSI scales used were: Arithmetic, Animal House, Mazes, Geometric Designs, and Block Designs. The latter four can be combined to compute a prorated performance I.Q. All children were tested individually according to the standard procedures. Testing usually took place in one forty-minute session. The WPPSI was administered during the same time intervals as the Stanford-Binet.

(3) Merrill-Palmer Scale of Mental Tests (MP). All sample children under four years of age were given the Merrill-Palmer. Verbal items in the Merrill-Palmer were not administered, but all other items were. This procedure allows for a computation of an I.Q. Children were tested individually following standard procedures. Testing usually lasted about one hour. The Merrill-Palmer was administered during the same time intervals as the Stanford-Binet and WPPSI. One tester gave all the Merrill-Palmer tests.

Measures of school-related behaviors and attitudes

(1) Birch procedure for categorizing response styles to cognitive task demands on the Stanford-Binet test. At the time of administration of the Stanford-Binet Intelligence Test, a record was made of the children's responses to the items presented and of the tester's behavior in conjunction with the test administration.

Hertzig, Birch, et al. (1968) have developed a coding system whereby children's responses to the Stanford-Binet can be examined for such dimensions as work orientation, type of not-work responses made, etc. This procedure originally was used in a study of lower-class Puerto Rican children and a contrast group of middle-class children. Ethnic and social-class differences were found in styles of response and it was hypothesized that these differences in part explained the differential school success of the groups under study. The procedure results in a number of scores on which group comparisons can be made. With minor modifications, we followed the Hertzig procedures, categorizing all responses made during the Stanford-Binet test.

(2) Testers' ratings of attention, task behavior, attitudes toward testers' authority. Following each administration of the Stanford-Binet, the examiner rated the child on a number of dimensions of test behavior. Ratings were made following both the first and second administration of the tests. The dimensions pertain broadly to test behavior as such and to the social relationship the child exhibits with the examiner.

These test ratings have been made every year in conjunction with the Ancona Headstart research. Test ratings were made on fifteen scales. The scales are designed so that end points do not necessarily conform to optimal vs. detrimental test behavior. End points of the scales usually refer to extremes of behavior with optimal behaviors falling at the midpoints. Ratings include such dimensions as Distractibility, Activity Level, Speed of Response, Persistence, Self-Confidence, and Understandability of Speech. A copy of the rating form can be found in Appendix A.

(3) Teachers' ratings of attention, task behavior, attitudes toward classroom authority. Teachers were asked to rate the children in the research sample on a number of dimensions at the beginning and end of the school year. Teachers' ratings were made at approximately the same time that the Stanford-Binet was administered. The rating forms used by the teachers were parallel to those used by the testers. The child's behavior in the classroom, rather than in the test situation, was the object under consideration. Twelve of the fifteen dimensions rated by the testers were rated by the teachers.

Measures of social interaction

(1) Social Interaction Observations. The procedure used for observing social interaction was a modification of the Marshall-McCandless (1957) Social Interaction method. A daily five-minute observation was made of each child until a total of 100 minutes of observation per child was collected. The goal of the observation was to record as closely as possible the child's social interaction and other activity behavior. The initials of each child with whom the observed child interacted, and instances of teacher-child interaction were noted.

Each record was coded by the observer according to the following method: the first social interaction occurring in each minute was coded, and if an act was directed to more than one child, the initials of each child involved were recorded. Thus a maximum of 100 acts per child were coded. Categories for coding social acts were:

I. Friendly interaction: associative play, friendly approach, friendly conversation, physical and verbal affection, nonverbal attention to initiation by others (e.g., smiling, nodding head, joining activity), imitation of others, compliance. Friendly interaction was further coded as follows:

a. dominative: child attempts to gain control of situation either diplomatically or authoritatively, giving directions, assigning roles, commanding.

b. integrative: interaction is cooperative with near equal sharing.

II. Aggression: Aggression was coded in the following manner:

a. physical: actual physical aggression or strong threat of same (e.g., raising hand, threatening with object).

b. verbal: name calling, threats, snubbing, withdrawal with verbal abuse, refusal to share or cooperate, criticizing, blaming, demanding.

c. provoked: aggression is in response to another child's threat.

d. unprovoked: aggression is initiated by the observed child.

III. Isolating Behavior: avoidance, withdrawal from activity without aggressive behavior (coded only if child does not join another group or immediate activity), ignoring the approach of another child (without aggressive behavior).

IV. Teacher-Student Interaction:

Recorded as T if initiated by teacher.

Recorded as S if initiated by student. Student initiated acts were further coded as SI if student seeks instrumental assistance, and as SE if student seeks emotional support, approval affection or comfort.

Social interaction observations were collected during the time period from late January through March, 1970. Four observers participated in data collection and coding. Reliability among the four observers paired with all partners for at least six simultaneous observation periods was calculated. The average agreement between all observers on both the occurance and category of an act was 93%. When agreement on to whom the act was directed was included, overall agreement was 85%. These reliabilities were considered satisfactory.

Operational Statement of Hypotheses for the Study of Nursery School Children

Having described the sample and instruments, it is now possible to detail the procedures for testing the hypotheses of the study.

Hypothesis 1: A Montessori program will increase the basic cognitive and behavioral skills of disadvantaged children required for educational achievement. This effect will be greatest if started early (age three) and continued over an extensive period of time.

Hypothesis 1a: The disadvantaged children will show increased cognitive development. We especially anticipate improvement in number concepts, visual discrimination, classificatory skills and psycho-motor skills which are emphasized in the curriculum.

Hypothesis 1b: The disadvantaged children will show increased attentiveness to task demands.

There are a number of analyses which should shed light on the validity of Hypothesis 1 and its subparts. In regard to long-term effects, it is possible to examine the performance of children who have been in the school more than one year to see if they improve intellectually in relation to their initial status. The data most relevant to this point are Stanford-Binet scores which are available at the beginning and end of successive years in school.

The effect of this year's participation on the cognitive development of the nursery school children will be examined by comparing mean performance on the Stanford-Binet at the beginning and end of the school year and by a similar analysis of the performance on the WPSSI and Merrill-Palmer tests. Comparisons on these tests can be made for the group as a whole as well as for children who are in their first year of school and children who have had previous schooling.

We are especially interested in examining the children's performance on the WPPSI and Merrill-Palmer tests. We chose the WPPSI scales because they seem to measure skills which we believe are emphasized in the Montessori curriculum. In contrast to the Stanford-Binet, the WPPSI yields standard scores on performance aspects of growth. It is in this area that we believe the Montessori curriculum should prove effective. Thus we will be particularly interested in contrasting the results on the WPPSI and Merrill-Palmer tests with those of the Stanford-Binet. Although the Merrill-Palmer does not yield separate scale scores, the content of the test items articulate fairly well with those of the WPPSI. Thus we chose the Merrill-Palmer for use with the young children in the sample for whom the WPPSI was inappropriate.

Assessment of children's performance in regard to school facilitative behaviors comes from a number of sources. Test and teacher ratings of behavior will be examined for possible changes from the beginning to the end of the year. In addition, the Birch procedure will be included in this analysis.

Hypothesis 2: Interaction between middle-class and disadvantaged children will increase as a function of experience in the program.

Hypothesis 2a: There will be more cross-group social acts in children with longer tenure in the school.

Hypothesis 2b: There will be more cross-group friendships among children with longer tenure in the school.

On the basis of the social interaction observations, we will compare the nature and quantity of social interactions exhibited by the various subgroups of our sample. It will be possible to compare the extent of friendly and aggressive interactions on the part of children in their first year of school and children with previous school experience, by social class and race. In addition, the observation procedure can be coded to produce a "best friend" score, based on each child's highest frequency of interaction with another child. These best friend scores can be used to assess hypothesis 2b.

Results and Discussion

Stanford-Binet Data

Stanford-Binet tests have been administered to the children at Ancona since the inception of the research program in the school. Last year we examined the performance of eleven children who were tested over a period of two years (September 1967 to June 1969) and found substantially no change in the mean Stanford-Binet scores over the two-year period of school attendance. Since the Stanford-Binet has only been administered at the nursery level, we do not have further data on most of the children in the group examined last year. However, fourteen children currently in the nursery group have been tested for two consecutive years. The mean performance on the Stanford-Binet for this group of fourteen children over two years is shown in Table 1.

Table 1

Mean Stanford-Binet I.Q. Scores at Four Time Points for Children with Two Years of Attendance at Ancona (N = 14)

Time of Test	Sept. 1968	June 1969	Sept. 1969	June 1970
Mean	86.43	94.36	91.64	92.57
Standard Dev.	10.41	10.95	8.71	13.30

An analysis of variance on these data results in an F ratio which is not significant at the usual statistical standards. As was the case last year, the performance of children with more than one year in the school does not support the hypothesis of continued intellectual acceleration. Consistent with other findings in the field is the initial increase in I.Q. followed by a slight decline or levelling. This pattern found by us, and noted by Bereiter (1966) and others seems to be quite general for disadvantaged preschool children. First school attendance does lead to some increase in intelligence test performance followed by substantially no change or decline to initial levels in subsequent years.

It is interesting to note that last year's group had an initial I.Q. level which was considerably higher than this year's group (96 vs. 86) but at the end of two years the performance of both groups is virtually equivalent (94 vs. 93).

We believe it appropriate on the basis of these data to conclude that the Montessori preschool experience does not lead to long-term increases in cognitive development in disadvantaged children when assessed by the Stanford-Binet. Initial school attendance does contribute to an increase in Stanford-Binet performance, but continuing increases are not to be expected from subsequent years in school. As we have argued previously (Stodolsky & Jensen, 1969), this finding may or may not portend the academic future of the children. Change on the Stanford-Binet is clearly not the only aspect of the child which would predict ultimate school adjustment and achievement.

We have further argued that the Stanford-Binet may not be the most appropriate assessment instrument to use in conjunction with the performance of children in a Montessori program. A number of major thrusts seem apparent in the Montessori curriculum if one examines the materials which the children use on a daily basis. Stress is on the development of matching and sorting skills, visual discrimination and visual-motor integration, eye-hand coordination, and the development of number concepts. We posit, therefore, that scales more directly relevant to such "performance" aspects of intelligence may show more growth in the sample than the Stanford-Binet. In this year's research, we have attempted to incorporate an assessment of such skills by administering the Merrill-Palmer and WPPSI scales to the children. After a presentation of the Stanford-Binet data for this year's sample we will turn to the evidence regarding outcomes on the Merrill-Palmer and WPPSI scales.

As a general orientation to the Stanford-Binet data for this year, Table 2 contains the means and standard deviations for the total sample and the social class groups at pre- and post-test times.

Table 2

Means and Standard Deviations for Stanford-Binet I.Q., M.A., and C.A. for the Lower-Class and Middle-Class Children Attending Ancona Preschool Classes 1969-70

	I.Q.		M.A.		C.A.	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Lower Class (N= 29)	90.76 (9.71)	94.28 (11.84)	44.69 (11.04)	51.72 (10.88)	48.31 (10.09)	54.45 (10.63)
Middle Class (N= 29)	116.28 (16.64)	122.00 (17.67)	58.41 (11.01)	68.07 (10.63)	50.14 (9.22)	56.38 (9.80)
Total (N= 58)	103.52 (18.65)	108.14 (20.44)	51.55 (12.94)	59.90 (13.48)	49.22 (9.63)	55.41 (10.18)

As is apparent in Table 2, the overall mean I.Q. for the Headstart sample is considerably below that for the middle-class comparison group. In a two-way analysis of variance, the social class effect on I.Q. produces a highly significant F ratio (p less than .01). Change over time from pre- to post-testing is not significant. Similarly, mental age differs significantly between the Headstart and middle-class group, whereas chronological age shows no difference.

From Table 2 it is apparent that both the lower-class and middle-class groups experienced a slight increase in I.Q. from the beginning to the end of the school year. However, these increases do not meet the usual standards of statistical significance. Generally we find the most apparent increases on Stanford-Binet I.Q. in children in their first year of school. Table 3 contains the Stanford-Binet scores of the sample children by social class and year in the school.

Table 3

Means and Standard Deviations for Stanford-Binet I.Q., M.A., and C.A. for the Ancona Nursery School Sample, by Year of Attendance and Social Class*

	I.Q.		M.A.		C.A.	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Lower-Class First Year	88.00 (10.11)	95.00 (10.67)	36.46 (9.19)	44.23 (8.11)	40.39 (6.86)	45.77 (6.86)
Middle-Class First Year	114.06 (14.86)	122.18 (18.55)	52.71 (9.37)	63.00 (9.23)	45.88 (8.34)	51.65 (8.75)
Lower-Class Years 2 to 3	93.00 (9.06)	93.69 (13.03)	51.37 (7.31)	57.81 (8.94)	54.75 (7.33)	61.50 (7.40)
Middle-Class Years 2 to 3	119.42 (19.10)	121.75 (17.14)	66.50 (7.70)	75.25 (8.25)	56.17 (6.90)	63.08 (7.03)

---

*N = 13 for Lower-Class First Year
N = 17 for Middle-Class First Year
N = 16 for Lower-Class Years 2 to 3
N = 12 for Middle-Class Years 2 to 3

As expected, the children in their first year in preschool do show the most gain on the Stanford-Binet. This increase is apparent in both the Headstart and middle-class children in their first year in the program. A two-way analysis of variance on these data does result in both a significant time and social-class effect for the first year children. Middle-class and Headstart children in their second and third years in the program do not experience change on the Stanford-Binet.

The data in Table 3 lend some support to hypothesis 1a. Disadvantaged children do improve on intelligence tests after their first school exposure. The fact that the middle-class first year children made similar gains is contrary to our findings of last year. Initial I.Q. levels of the middle-class samples from this year and last year are virtually identical. The reason for gains this year is not apparent.

Taking the data on cumulative Stanford-Binet performance shown in Table 1 together with the data on the sample in the nursery school this year, a cautious conclusion about the effect of preschool experience on cognitive development seems warranted. We have rather consistently found an initial increase in I.Q. for children in their first year of school experience: this is uniformly the case for disadvantaged children and sometimes the case for advantaged children. After the first year of school, no further changes in intelligence test performance have been found. Workers in many intervention projects have had similar findings and a number of possible explanations have been put forward. It may well be that gains after first school exposure are due more to rapport and test effects (Zigler and Butterfield, 1968) than to cognitive growth. Alternatively, change may in fact reflect growth but the curriculum is not sufficiently powerful to do more than has been accomplished in the first year of schooling to boost intelligence test performance. Thus children in preschool programs do experience gains and the gains are maintained while the children are still in the program (see Grottberg, 1969). A third possibility is that there are limitations within the child himself which place a ceiling on the amount of change to be expected in general intelligence. Thus nurture can enhance the child's performance up to a point, but biological limits still operate as a limit on the process.

Regardless of the proper explanation or combination of explanations for these data, it is still very clear that the ultimate concern must lie with the child's academic achievement and eventually his functioning as an adult in the society. Furthermore, from a curricular point of view, evaluation must be tied closely to the processes which are emphasized in the curriculum. The ubiquitous use of general intelligence tests has probably severely limited our understanding of curricular variations and their differential impact.

In an attempt to come closer to a sensitive evaluation of the Montessori curriculum, we administered Merrill-Palmer and WPPSI scales to the nursery children at Ancona. The former was administered to young children (under age four) and the latter to children age four and older. Results for children given the WPPSI are presented in Table 4 for the two social class groups.

Table 4

Means and Standard Deviations of WPPSI I.Q. and Performance Scales, for the Ancona Nursery Sample, by Social Class (N= 47)

	Performance I.Q.				Arithmetic				Animal House			
	Time 1	Time 2	t^1	p	Time 1	Time 2	t	p	Time 1	Time 2	t	p
Lower-Class (N= 21)	87.67	94.71	4.49	001	7.43	9.00	3.83	005	7.71	9.38	3.79	005
	(9.13)	(9.74)			(2.36)	(2.43)			(2.49)	(1.94)		
Middle-Class (N= 26)	101.27	114.31	4.66	001	10.11	11.69	2.98	01	10.08	11.81	2.40	025
	(18.81)	(18.59)			(3.72)	(2.80)			(3.70)	(3.70)		
Total	95.19	105.55	5.34	001	8.91	10.49	4.28	001	9.02	10.72	3.54	001
	(16.59)	(18.05)			(3.43)	(2.94)			(3.40)	(3.25)		

	Mazes				Geometric Design				Block Design			
	Time 1	Time 2	t	p	Time 1	Time 2	t	p	Time 1	Time 2	t	p
Lower-Class	8.24	8.62	1.41	N.S.	8.43	9.57	1.63	N.S.	8.67	10.48	3.93	001
	(2.14)	(2.13)			(2.69)	(3.31)			(2.50)	(1.83)		
Middle-Class	10.04	12.23	5.21	001	9.27	11.62	3.31	005	11.19	13.31	3.53	005
	(3.73)	(3.76)			(3.85)	(2.83)			(3.20)	(3.27)		
Total	9.23	10.62	4.79	001	8.89	10.70	3.62	001	10.06	12.04	4.71	001
	(3.22)	(3.60)			(3.38)	(3.19)			(3.14)	(3.05)		

---

[1] t computed using paired-comparison method.

As is evident from Table 4, the children given the WPPSI do show considerable change on this instrument in marked contrast to their performance on the Stanford-Binet. Again we find marked social class differences which are statistically significant, but in this case time of testing also produces a significant effect in a two way analysis of variance. Prorated Performance I.Q., Arithmetic, Animal House, and Block Design show significant changes for each of the social-class groups and for the combined sample. On Geometric Design and Mazes, only the middle-class children show a significant gain. We were particularly interested in administering this instrument because if it is sensitive to the emphases of the Montessori curriculum we would expect change on the part of both the disadvantaged and middle-class children. We believe the data in Table 4 generally indicate a systematic program effect in evidence for the children.

We selected the specific scales of the WPPSI on a number of bases. The Arithmetic scale was chosen because of the stress on number concepts in the Montessori math curriculum and the considerable number of materials available for the development of arithmetic skills such as the various number bead exercises. It should be noted that the Arithmetic scale is not used in calculating prorated Performance I.Q., but rather is a subscale in the Verbal I.Q. on the WPPSI. Animal House is a visual matching task; the curriculum provides many experiences in matching and sorting for the children. Block Design is the copying of designs made with blocks. Most items are presented with a model in front of the child, the most difficult items are a picture of a design to be made from blocks. Wechsler considers Block Design to be

"... a sorting as well as a perceptual motor task." (WPPSI Manual, 1967). We believed the children's experience with a variety of block and puzzle games, with sorting tasks, and with geometric insets would facilitate their performance on Block Design. Geometric Design involves drawing geometric figures from a model. We believed this scale would measure the children's skills in eye-hand coordination and the use of the pencil, as well as drawing upon the many materials which are geometric designs such as drawing insets. We gave the Mazes scale in order to have enough scales to compute a performance I.Q. and because we believed it would measure eye-hand coordination and use of the pencil. In addition, we recognized that Mazes does tap spatial visualization skills, an area which is not necessarily emphasized in the Montessori curriculum.

Evidently the curriculum as experienced by the lower-class children was not as effective as we hypothesized in developing fine motor, use of pencil skills, as witnessed by their non-significant change on the Mazes and Geometric Design scales. It may be that although these scales tap eye-hand coordination, they go beyond that to measure such abilities as spatial visualization and directional concepts which are not well-developed in the disadvantaged children. It might also be that the children are not utilizing the particular materials available in these areas, but through self-selection of activities are concentrating elsewhere (See Karlson, 1970, in preparation).

It should be recalled that most of the children who received the WPPSI were above four years of age. The Manual specifies that the lower age limit should be close to four years, but that bright children below this age will be testable. In general, we only gave the WPPSI to children below four years if they seemed comfortable in the test situation and able to perform. The decision was based on examiner's judgments of the children. As noted in our description of the sample, more of the middle-class first-year children were closer to four than the lower-class first-year children. Taking both chronological age and social class into account, resulted in the fact that many more first-year middle-class children received the WPPSI than did first-year lower-class children. Thus, most of the lower-class children in Table 4 are second to third year in school, whereas the middle-class group shown in Table 4 is a more even mixture of children from first through third year in school.

Of special interest to us is whether change on the WPPSI will be evident in children who are beyond their first year in school. We have noted that such change is most unusual on the Stanford-Binet, and have argued that this may be the case because the Binet does not articulate well with the curricular emphases of the school. If our contention is correct that the WPPSI is a closer measure of school objectives, change should be evident for children beyond their first year in school. Also, we are sure that the measure is valid for the children in the older group, whereas some of the younger children may show change spuriously as the initial estimates may have been too low if the test was in fact too difficult for them. Table 5 contains the WPSSI means and standard deviations for the sample, split by social class and year in the program.

Table 5

Means and Standard Deviations of WPPSI I.Q. and Performance Scales, for the Ancona Nursery Sample, by Social Class, and Year in Program (N = 47)

	Performance I.Q.				Arithmetic				Animal House			
	Time 1	Time 2	t[1]	p	Time 1	Time 2	t	p	Time 1	Time 2	t	p
LC Year 1 (N=5)	82.60	98.60	6.23	<005	8.00	8.80	.776	N.S.	8.00	10.40	2.42	<10
Mean Age = 45.6	(8.88)	(7.40)			(1.58)	(3.11)			(2.55)	(1.67)		
LC Year 2 (N=16)	89.25	93.5	2.28	<05	7.25	9.06	3.85	<005	7.63	9.06	2.51	<025
Mean Age = 54.8	(8.89)	(10.26)			(2.57)	(2.29)			(2.55)	(1.95)		
MC Year 1 (N=14)	92.86	107.64	3.75	<005	8.43	11.14	2.79	<025	8.86	10.71	1.73	N.S.
Mean Age = 53.4	(18.69)	(19.78)			(3.65)	(2.98)			(3.90)	(4.16)		
MC Year 2 (N=12)	111.08	122.08	3.55	<005	12.08	12.33	.32	N.S.	11.50	13.08	1.93	<10
Mean Age = 56.3	(14.01)	(14.16)			(2.81)	(2.54)			(3.00)	(2.71)		

	Mazes				Geometric Design				Block Design			
	Time 1	Time 2	t	p	Time 1	Time 2	t	p	Time 1	Time 2	t	p
LC Year 1 (N=5)	7.60	8.60	1.60	N.S.	6.60	8.80	1.52	N.S.	7.40	11.20	3.92	<025
Mean Age = 45.6	(2.51)	(1.14)			(2.61)	(3.42)			(2.61)	(1.30)		
LC Year 2 (N=16)	8.44	8.63	.19	N.S.	9.00	9.81	1.02	N.S.	9.06	10.25	2.29	<05
Mean Age = 54.8	(2.06)	(2.39)			(2.53)	(3.35)			(2.41)	(1.95)		
MC Year 1 (N=14)	8.71	11.14	5.03	<001	8.21	10.86	3.01	<01	9.79	12.50	3.04	<01
Mean Age = 53.4	(3.67)	(4.06)			(3.40)	(2.77)			(3.17)	(3.84)		
MC Year 2 (N=12)	11.58	13.50	3.62	<005	10.50	12.50	1.82	<10	12.83	14.25	2.22	<05
Mean Age = 56.3	(3.29)	(3.06)			(4.12)	(2.75)			(2.44)	(2.26)		

As can be seen from Table 5, all groups show significant change on the prorated performance I.Q. from the beginning to the end of the school year. The disadvantaged children in their second and third years of preschool change significantly on Arithmetic, Animal House and Block Design the same result which was obtained for the combined lower-class sample. The middle-class first-year group shows significant change on all scales with the exception of Animal House which is near significant. The lack of change on this scale seems to be accounted for by extreme variability in the performance of the group on this sorting task. The second-year middle-class group shows significant change on Mazes and Block Design. The change for this group on Geometric Design and Animal House approaches significance. The older middle-class children clearly do not experience change on the Arithmetic scale. The reason for this may lie in the nature of the Arithmetic scale. Raw scores for this group indicate mastery of all the items which are concrete (presented with blocks, etc.). Difficulty arises when the items are read to the child and are word problems. The Montessori materials do not really teach such skills at this level. The younger children and lower-class children improve because their initial achievement is below this cutting point in the scale.

Taken together, the results shown in Table 5 seem to support a program effect which is demonstrable using the WPPSI. This is clearly the case for the disadvantaged children in later years of preschool who do show change in some of these performance measures while not changing on the Stanford-Binet. The middle-class first-year children seem to have shown growth on both the

WPPSI and the Stanford-Binet and the middle-class older children have shown some change on the WPPSI, but it is not as dramatic as for the lower-class children.

Considering the fact that the WPPSI is only a rough attempt to measure the specific objectives of a Montessori curriculum and is certainly not tailor-made for the situation, these data suggest the fruitfulness of further evaluation attempts along these lines.

Research on the effects of the Montessori curriculum has been rather limited, but a number of workers were similarly interested in verbal and performance outcomes. (See Berger, 1968 as illustrative.) To our knowledge, the WPSSI has not been administered in any previous studies. However, Bereiter (1967) administered the ITPA and Wide Range Achievement tests to a group of middle-class children. In a comparison of the Montessori curriculum and the Direct Verbal Instruction method, he found few differences. He concluded:

> The differences between groups are minute and contrary to the expectation that the contrasting verbal and nonverbal treatments given the two groups would be reflected positively in ITPA subtest changes. ... It appears from the present results that Montessori trained children do not do better on non-verbal tasks *per se*, but only on non-verbal tasks that have a minimum of conceptual content.

It should be recalled that his sample consisted of bright middle-class children. It may also be the case that the ITPA does not tap the most salient aspects of non-verbal performance developed in the Montessori curriculum.

Berger (1969) did an extensive comparison of Montessori and traditionally trained pre-school disadvantaged children. She hypothesized as we have that the Montessori trained children would show improvement in the perceptual area, especially visual-analytic and visual-motor integration skills. Generally, her findings for four-and five-year olds tend to support her hypothesis. She says,

> With respect to perceptual training trends, the public school and community center studies concurred. Results in each case consistently discriminated in favor of the Montessori trained children, in support of the hypothesis, and at both age levels investigated. (Berger, 1969, p. 49)

Of the perceptual tests she used, the Montessori children showed superior performance on a Block Design task, and some groups showed superior performance on Mannikin (from the WISC), and a puzzle. Of interest is her finding that the low achievers seemed to benefit the most from the perceptual training emphasis. Differences in favor of the Montessori trained children in the perceptual area were still present after two years of schooling as evidenced by certain scales on the Frostig Battery. In contrast to the perceptual area, cognitive tasks did not show a systematic trend in favor of the Montessori group. Results were mixed, but where differences pertained they tended to favor the traditional kindergarten program. Berger's findings and ours tend to be consistent. On the basis of our findings in the nursery sample and similar findings for the WISC with the older children in the elementary program (See Part II), we believe this line of inquiry should be followed further.

Since the WPPSI was not age-appropriate for all of the children in the Ancona sample, the Merrill-Palmer was administered to children younger than four years of age in an attempt to assess similar performance aspects of the children's growth. The results of the administration of the Merrill-Palmer are shown in Table 6. A total of twenty-three young children were given the Merrill-Palmer scale. The average age at first testing for the children was approximately forty months. There was no age difference between the lower and middle-class children who received the Merrill-Palmer.

Table 6

Means and Standard Deviations for Merrill-Palmer I.Q., M.A., and Percentile Rank for the Young Children at Ancona Nursery by Social Class (N=23)

	I.Q.		M.A.		Percentile Rank	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Lower Class	108.42	118.67	44.25	54.58	69.08	81.33
(N=12)	(11.84)	(14.46)	(8.80)	(8.98)	(18.90)	(11.44)
Middle Class	129.45	134.00	51.36	61.09	86.82	94.00
(N=11)	(22.27)	(16.28)	(8.88)	(7.03)	(16.86)	(7.48)
Total	118.48	126.00	47.65	57.70	77.57	87.65
	(20.27)	(16.92)	(9.37)	(8.59)	(19.75)	(11.38)

As can be seen in Table 6, the Merrill-Palmer performance of these children did show improvement from the beginning to the end of the year. The change for the lower-class children and the total sample is statistically significant beyond .01. The middle-class young children showed some improvement, but the change is not statistically significant. These children are the bulk of the first year group we examined in relation to the WPPSI where they did exhibit significant change. The lack of change on the Merrill-Palmer seems due to ceiling effects in this test. Their initial level is very high and the scale does not seem to have enough difficult items for them to show much improvement. This fact is supported by their post-test percentile rank of 94 accompanied by a great decrease in the standard deviation at post-test time.

We administered almost the entire Merrill-Palmer scale to the children, eliminating the vocabulary scale. The test does not provide for subtest scores, but we have examined the performance of the children on various aspects of the test to see if some areas show more improvement than others. Unfortunately the test does not contain any scales pertaining to number or arithmetic. However, the lower-class children show most improvement on items which are puzzles or puzzle-like (Mare and Foal, Picture Puzzles and Form Board), and on Matching, Peg Board, Buttoning, and Geometric Designs. The transfer from the curriculum seems apparent in considering these tests. Least improvement is apparent on the Pink Tower, Pyramids, and Mannikin items. The first two are both block building items and Mannikin is a body-part puzzle.

As with the WPPSI, we feel that the Merrill-Palmer was tapping more of the skills which develop through the Montessori experience. The over-all level of performance of the lower-class children on this test is very high, but the norms are old and therefore caution is necessary in interpreting the absolute figures.

All test data which pertains to Hypothesis 1a has now been presented and discussed. In summary, it seems apparent from these data that the Montessori curriculum is more powerful in the performance than verbal areas. Generally, we have found change on the Stanford-Binet in children during their first year of attendance at Ancona. After the first year of schooling change is not usually in evidence and continuing children level off or show slight declines in Stanford-Binet I.Q. after their first year of school. Our findings on the WPPSI and Merrill-Palmer support the contention that the Montessori curriculum is more effective in the performance areas. We have seen change on these measures for children in their first year of school and for children in subsequent years.

We would predict that a test battery which contained an even broader spectrum of measures which tap the curricular emphases would show an even clearer picture of this sort. Unfortunately, we do not know whether the performance emphasis of the curriculum will be facilitative of school success for these children in the long run. On the basis of the follow-up data now available to us, and the acknowledged verbal emphasis of most later schooling, we do not anticipate much pay-off for the children. One then must ask whether the program in the early years should be supplemented with a more language-oriented component or if it is possible to restructure elementary level instruction to capitalize on the performance area strengths of the Montessori nursery graduates.

Having discussed the intellectual measures for the sample, we will now turn to Hypothesis 1b which deals with other behaviors which may be facilitative of school success. In this connection, teacher and tester ratings of behavior and the behavior analyzed using the Hertzig-Birch procedure are of interest.

Test and Teacher Ratings

In past years' research at Ancona, particularly 1965-66, a number of characteristics rated by the testers and teachers were found to change significantly over the year for Headstart children (Jensen and Kohlberg, 1966). The ratings which deal primarily with task orientation showed the most change and some showed a correlation with intelligence test change scores. In particular, Distractibility was found to be very high among the entering Headstart children and to show a noticeable decline, particularly for children who showed an increase in I.Q. on the Stanford-Binet.

Our findings last year (Stodolsky and Jensen, 1969) were in marked contrast to the previous findings at Ancona. The most salient aspects of our findings last year were: 1) the middle-class comparison group and the Headstart children were highly similar on test and teacher ratings; 2) the first-year children were most similar; 3) ratings were generally uncorrelated with performance on the Stanford-Binet; and 4) few changes occurred on the ratings from the beginning to the end of the school year.

We have performed a number of analyses on test and teacher ratings in order to assess the initial status of our group and possible changes which would occur as a result of attendance at Ancona. In particular, we have compared the Headstart children with the middle-class comparison group at both time points to see if the groups differ. Next, we have examined the groups separately at Time 1 and Time 2 to see if there is change over time. In addition, we have correlated test and teacher ratings with each other and with intelligence test performance.

It should be recalled that the testers' ratings were made in conjunction with the administration of the Stanford-Binet. This year all children were tested by the same examiner (one of three) at both times so that the ratings for each child were made by the same person at Time 1 and Time 2. Teacher ratings of classroom behavior were made at approximately the same time that the intelligence tests were administered.

The testers rated fifteen characteristics of the test behavior of the children during the administration of the Stanford-Binet. The teachers rated twelve characteristics of the children with respect to their classroom behavior. The scales can be categorized broadly into two domains: characteristics dealing with task orientation and characteristics dealing with social relationships, particularly relationships to adults. In the early research at Ancona, the ratings dealing with task orientation have shown the most change and have been considered most important in terms of facilitating intellectual change. The few changes found last year were also on ratings in the task orientation area.

The ratings which fall under the general category of task orientation are: Distractibility, Activity Level, Speed of Response-Verbal, Speed of Response-Performance, Initiative with Materials, Self-Confidence on Tasks, Persistence, Reaction to Failure, Sense of Intellectual Challenge, and Willingness to Continue. The scales dealing with social relationships to the adult are: Fear of Adult, Social Initiative-Adult, Communication of Affect, and Compliance with Adult. The Verbalization scale which refers to understandability of speech cannot be placed in either of these general categories.

Table 7 contains the means and standard deviations for the

Table 7

Mean Ratings by Testers and Teachers for the Ancona Sample, by Social Class, Length of Attendance and Time

	Lower-Class Total (N = 29)*				Middle-Class Total (N = 29)*				Lower-Class First Year (N = 13)*			
	Time 1		Time 2		Time 1		Time 2		Time 1		Time 2	
Test Ratings	M.	S.D.	M.	S.D.	M.	S.D.	M.	S.D.	M.	S.D.	M.	S.D.
Distractibility	5.7	2.1	5.4	2.1	4.3	1.7	4.1	1.9	6.2	2.5	6.0	2.
Activity Level	4.5	2.0	3.7	2.0	4.8	1.5	4.8	1.9	4.0	2.4	3.5	2.
Speed of Response-Verbal	5.0	2.0	5.6	1.9	4.2	1.6	4.5	1.7	6.2	2.2	5.7	1.
Speed of Response-Performance	4.0	1.6	3.6	1.5	3.6	1.4	3.6	1.6	4.4	2.0	4.0	1.
Initiative-Test Materials	3.9	2.1	3.7	2.2	4.3	1.7	4.7	1.7	4.0	2.6	3.1	2.
Self-Confidence on Tasks	4.5	2.4	5.2	2.1	4.8	1.6	5.0	1.5	4.4	2.6	5.2	2.
Persistence	5.7	1.9	6.1	1.6	5.0	1.5	4.3	1.2	5.4	2.4	5.9	1.
Reaction to Failure	2.0	1.9	2.2	2.2	2.5	1.8	2.8	1.6	2.5	2.2	2.5	2.
Sense of Intellectual Challenge	6.0	1.8	6.2	1.3	4.7	1.5	4.3	1.5	6.5	2.2	6.1	1.
Willingness to Continue	5.5	2.1	5.8	2.2	4.7	2.0	4.6	2.1	5.7	2.3	5.9	2.
Fear of Adult	4.6	1.8	3.7	2.2	3.4	1.9	3.4	1.9	4.5	2.1	3.8	2.
Social Initiative with Adult	4.6	2.3	4.5	2.0	4.5	1.6	4.4	1.6	4.1	2.8	4.1	2.
Communication of Affect	3.6	1.9	3.9	2.1	4.1	1.5	4.1	1.7	3.0	1.7	3.4	2.
Compliance with Adult	5.2	1.8	5.5	1.5	4.5	1.4	4.5	1.1	5.5	2.1	5.8	1.
Verbalization	5.4	2.3	4.9	2.4	2.7	2.2	2.1	1.6	5.9	2.4	6.4	2.
Teacher Ratings												
Distractibility	5.3	1.9	5.3	1.9	4.9	1.7	4.9	1.8	5.3	2.6	5.3	2.
Activity Level	3.9	1.8	3.7	1.6	3.7	1.6	3.7	1.3	3.4	1.8	3.1	1.
Initiative-Classroom Materials	3.8	2.0	4.7	1.8	3.8	1.7	4.3	1.8	3.3	2.7	4.7	2.
Self-Confidence on Tasks	5.1	2.0	5.1	2.0	5.1	1.7	4.5	1.8	4.3	2.4	5.2	2.
Persistence	5.3	1.4	5.6	1.5	5.5	1.2	5.0	1.0	5.9	0.8	6.0	2.
Reaction to Failure	3.9	2.1	4.0	2.7	4.4	2.0	5.0	1.6	3.0	2.3	3.9	3.
Sense of Intellectual Challenge	5.5	2.3	5.8	2.0	5.1	1.7	4.9	1.7	5.8	3.0	6.4	2.
Fear of Adult	4.1	2.1	3.7	2.7	3.5	2.0	3.1	1.7	4.4	2.6	4.3	3.
Social Initiative with Adult	4.4	1.4	4.3	1.5	4.0	1.2	3.5	1.2	4.9	1.9	4.8	1.
Communication of Affect	3.7	1.8	4.1	1.8	3.5	1.6	3.1	1.3	3.5	2.4	4.5	2.
Compliance with Adult	5.5	1.3	5.7	1.0	5.1	1.1	5.2	1.9	5.9	1.2	6.1	1.
Verbalization	5.2	2.6	4.3	2.7	2.7	2.3	2.3	1.8	6.4	3.2	5.5	3.

*N's for testers' ratings are as shown. For teachers, LC total = 24, MC total = 27, LC Year 1 = 8, MC
2 = 15. Variation due to late entrance of some children.

Middle-Class First Year (N = 17)*				Lower-Class 2 to 3 Years (N = 16)				Middle-Class 2 to 3 Years (N = 12)				
Time 1		Time 2		Time 1		Time 2		Time 1		Time 2		
M.	S.D.	M.	S.D.	M.	S.D.	M.	S.D.	M.	S.D.	M.	S.D.	**Test Ratings**
4.4	1.8	4.4	2.1	5.3	1.8	4.9	2.2	4.2	1.7	3.8	1.6	Distractibility
5.0	1.6	5.1	2.0	4.8	1.7	3.9	1.8	4.5	1.3	4.4	1.7	Activity Level
4.4	1.6	4.6	1.7	4.1	1.2	5.5	2.0	4.1	1.7	4.3	1.9	Speed of Response-Verbal
3.7	1.4	3.5	1.6	3.6	1.3	3.3	1.0	3.5	1.4	3.7	1.7	Speed of Response-Performance
4.2	1.9	5.1	1.5	3.9	1.7	4.3	2.0	4.4	1.4	4.1	1.7	Initiative-Test Materials
4.8	1.6	5.1	1.2	4.5	2.2	5.2	1.8	4.7	1.5	4.8	1.9	Self-Confidence on Tasks
5.0	1.5	4.6	1.3	6.0	1.4	6.3	1.4	5.0	1.6	3.8	0.9	Persistence
2.5	1.7	2.6	1.8	1.6	1.6	2.1	1.8	2.5	2.0	3.1	1.2	Reaction to Failure
5.0	1.5	4.8	1.5	5.7	1.3	6.3	1.1	4.2	1.5	3.6	1.4	Sense of Intellectual Challenge
5.2	2.1	4.9	2.5	5.3	1.9	5.8	2.3	4.0	1.5	4.2	1.6	Willingness to Continue
3.7	2.0	3.6	2.1	4.7	1.5	3.5	2.1	3.0	1.6	3.3	1.8	Fear of Adult
4.7	1.6	4.5	1.9	5.1	1.8	4.8	1.6	4.2	1.5	4.3	1.1	Social Initiative with Adult
4.5	1.5	4.3	1.7	4.1	2.0	4.4	1.8	3.6	1.5	3.8	1.7	Communication of Affect
4.5	1.6	4.5	1.3	4.9	1.6	5.3	1.4	4.5	1.2	4.7	1.0	Compliance with Adult
3.6	2.5	2.5	1.9	5.1	2.2	3.6	1.7	1.5	0.7	1.6	0.8	Verbalization
												Teacher Ratings
4.7	1.9	5.2	2.0	5.4	1.5	5.3	1.5	5.3	1.5	4.5	1.6	Distractibility
4.0	1.7	3.6	1.5	4.2	1.8	4.1	1.6	3.3	1.4	3.9	1.1	Activity Level
4.1	1.9	4.2	1.9	4.1	1.6	4.7	1.7	3.5	1.2	4.3	1.6	Initiative-Classroom Materials
5.0	1.9	4.5	1.9	5.6	1.6	5.1	2.0	5.3	1.4	4.7	1.6	Self-Confidence on Tasks
5.8	1.1	5.3	1.0	5.0	1.5	5.4	0.9	5.1	1.2	4.7	1.9	Persistence
4.6	2.4	5.0	1.5	4.4	2.0	4.1	2.4	4.3	1.5	4.9	1.7	Reaction to Failure
5.6	1.3	5.2	1.9	5.3	1.9	5.5	1.8	4.4	2.0	4.6	1.5	Sense of Intellectual Challenge
4.2	1.8	3.2	1.7	3.9	1.9	3.4	2.5	2.7	1.9	3.1	1.8	Fear of Adult
4.1	1.2	3.7	1.2	4.2	1.0	4.1	1.4	3.9	1.3	3.3	1.1	Social Initiative with Adult
3.8	1.8	3.5	1.6	3.8	1.5	3.9	1.4	3.2	1.4	2.7	0.7	Communication of Affect
3.2	1.3	3.5	0.7	5.3	1.3	5.4	0.7	5.0	1.7	4.9	1.0	Compliance with Adult
3.2	2.8	2.8	2.2	4.6	2.2	3.6	2.2	2.0	1.4	1.6	0.8	Verbalization

teacher and tester ratings for the entire lower- and middle-class samples and for these groups divided by year of attendance. For the total middle- and lower-class groups at Time 1, four scales show significant differences. Both the testers' and teachers' ratings of the lower-class children on Verbalization place them below the middle-class sample on understandability of speech. The speech of the lower-class sample is characterized as "adequate, there are errors, but speech is still easily understood." On all other teacher ratings there are no significant differences between the groups at Time 1. The testers' ratings yield three other significant differences. Testers rated the lower-class children as significantly more distractible; the level of the rating indicates normal attentiveness with some tendency for the children's attention to wander. The testers' ratings on Sense of Intellectual Challenge characterized the lower-class children as expending appropriate levels of effort on problems, but possibly more comfortable with easy tasks. The last difference at Time 1 was on the Fear of Adult scale. The Headstart children were rated as "neither confident nor fearful" whereas the middle-class children were "rather confident."

When the ratings are examined by year in program, a few more differences emerge. The first-year Headstart children differ from the middle-class comparison group on five testers' ratings and one teacher rating. Again, both testers' and teachers' ratings of these children differ on Verbalization. In addition, testers rated these children as more Distractible, slower in Response to Verbal Items, lower on Sense of Intellectual Challenge (but neither group exhibits a high level of challenge), and more Expressive of Affect (but with self-control).

These differences at Time 1 for the first-year children lie mostly in the area of task orientation measures and are not overwhelmingly. The entering children are more similar to one another on most scales than different with the clear exception of Verbalization, a finding which has replicated every year in our research. The fact that most differences were in tester ratings rather than teacher ratings is not easily explained.

Similarly, the Headstart children in their second and third years in school at Time 1 were rated more poorly on Verbalization than the middle-class group by the testers and teachers. They are also significantly different from their middle-class counterparts on the testers' ratings of Sense of Intellectual Challenge (less challenged by hard tasks), and Fear of Adult (neither confident nor fearful versus rather confident). Otherwise the two groups are highly similar on tester and teacher ratings.

Generally, then, at the beginning of the school year the Headstart and middle-class children were seen as highly similar in most respects. Verbalization ratings are a marked exception to this, and such differences have been found repeatedly in our research. Teacher ratings show fewer differences than tester ratings; and first-year children show more differences than the older children. (It should be recalled that the last year's first-year group showed few differences between the social classes and these are most of this year's older group.)

Having looked at the initial status of the children as reflected in the ratings, the next question of interest is whether these ratings of the children change from the beginning to the end of the school year. For the total sample, few changes meet the usual levels of statistical significance. The lower-class total group changes on the testers' Fear of Adult scale, becoming more confident. The middle-class total sample changes on the testers' rating of Persistence, becoming more persistent. Otherwise the group means from Time 1 to Time 2 are quite similar.

When the children's ratings are examined by year in program, both the lower-class and middle-class first-year children show no significant changes on the rating scales. Some changes are evident in the older children. The older Headstart children are rated somewhat slower to respond to verbal items at the end of the year, and the testers' ratings of their speech show marked improvement. Teacher ratings of Verbalization for this group also show improvement, but the change does not reach a statistically significant level. Middle-class older children change on the Persistence scale in the direction of more persistence. Although few changes occur for each group in the sample which are statistically significant, an inspection of social-class differences at Time 2 reveals more differences between the groups than were extant at Time 1. Thus shifts within each group have occurred so that they are less similar to one another at Time 2.

In regard to the first-year children, both teachers' and testers' ratings on Verbalization still show significant social-class differences; the mean ratings for both groups show improvement but the relative superiority of the middle-class children is still present. No other teachers' ratings for this group show social-class differences. There are social-class differences in the testers' ratings at Time 2. The lower-class children are still seen as more Distractible and experiencing less Sense of Intellectual Challenge than the comparison youngsters at Time 2. In addition, testers place the lower-class children as more active, as more assertive (perhaps unrealistically) on the Self-Confidence scale, as more upset by failure on the Reaction to Failure scale, and as less compliant with adults. The levels of these ratings are not extreme, but the differences generally would be characterized as favoring the middle-class children. It is interesting that most of these changes are due to a shift in the rating of the lower-class children and not to a shift in the middle-class children, although this is not uniformly the case.

The older children at Time 2 again show social-class differences in Verbalization ratings by teachers and testers. This is the case even though the lower-class children did improve significantly in Verbalization. The lower-class children continue to be rated lower on Sense of Intellectual Challenge at Time 2 by the testers. In addition, both teachers and testers place the lower-class children as less persistent than the middle-class children, but this difference is accounted for by the growth in this area on the part of the middle-class children and not by deterioration in the Headstart children.

In addition, the testers' rating on Willingness to Continue show a social-class difference, but the level of ratings for both groups on this scale seem acceptable. The teachers characterize the Headstart children as somewhat more expressive of affect than the comparison children, again the level of ratings for the groups seem within acceptable limits.

We take our general findings on the ratings to indicate a fairly high degree of similarity between the children in the Ancona classrooms. Somewhat more differences by social class are evident in the first-year children, especially by the end of the year than in the older children. Overall, these ratings do not seem to indicate much change in school-related behaviors. This may be a function of the fact that the initial levels for the children generally are within the desirable limits of the scales.

It is interesting that both this year and last year, there were more similarities on these scales than differences. Last year we found little relationship between the ratings and intelligence test performance, a finding which raised questions about the validity of the ratings as predictive of cognitive functioning. This year we have also attempted to examine the consistency of the ratings between teachers and testers and their relationship to the I.Q. measures.

In regard to ratings by teachers and testers, we find some scales do show consistency across these groups. In particular, Verbalization ratings are quite consistent when comparing teacher and tester judgments (r=.625 at Time 1, .579 at Time 2) and Social Initiative with Adults also shows significant positive correlations at Time 1 and Time 2. There is also an indication of consistency on the ratings of Sense of Intellectual Challenge, and

Communication of Affect (at Time 1 only). Activity level shows a significant negative correlation at Time 1 (r=-.321) and a significant positive correlation at Time 2 (r=.428). On other scales the correlations between teacher and tester ratings is near zero. It is interesting that there is more agreement in the area of social relationships than there is on the task orientation scales.

In our previous work we found little relationship between the ratings and Stanford-Binet I.Q. level when examined for each social class group separately. Currently, we have examined the correlations between ratings and Stanford-Binet I.Q. for our total sample. At Time 1, four of the fifteen tester ratings show a significant correlation with I.Q. level. These scales are Distractibility, Sense of Intellectual Challenge, Fear of Adult, and Verbalization. The teachers' ratings of Verbalization and Social Initiative with Adults also show a correlation with I.Q.

At Time 2, testers' ratings of Persistence, Sense of Intellectual Challenge, Compliance and Verbalization show significant correlations with I.Q. Teachers' ratings of Persistence, Sense of Intellectual Challenge, Initiative with Classroom Materials, Fear of Adult, Social Initiative with Teacher, Communication of Affect and Verbalization all show significant correlations with Time 2 Binet I.Q. These results lend some support to the contention that at least some of these rating scales are indicators of behaviors related to cognitive achievements. It is clear that the scales on which both the teachers and testers are consistent, are also the more likely to correlate with intelligence test performance. Thus we see such relationships operating relative to

Persistence, Sense of Intellectual Challenge, Verbalization, and Social Initiative with Adults.

Although there is reason to assume some validity in these scales, the finding of little change on them tends to suggest few changes in the children on these indicators of school-facilitative behaviors. To the extent that this is a valid picture, it appears primarily to be a function of high initial similarity among the children. In addition to the ratings, we have examined the children's behavior during the administration of the Stanford-Binet, using the measures of responsiveness to cognitive demands developed by Hertzig and Birch. We turn to these data now to see whether we can find consistency in this behavioral area.

Birch Procedure

The Birch procedure is another set of measures which assesses tack orientation in children. Last year we used a modification of the original system developed by the University of Hawaii Headstart Evaluation and Research Center. The main feature of that modification was that it utilized only the final response given by the child to each item in the Stanford-Binet administration. This year, we have abandoned the Hawaii modification, and gone back to the original procedure developed by Hertzig, Birch, Thomas and Mendez (1968) with only minor modifications. This year's data then are based on all the responses made by the children and not just the final response.

The procedure was used originally by its authors, and by us, to categorize children's responses to the task demands on the Stanford-Binet test. The system of categories is described by the authors as arising from the objective possibilities for responding, expressed as a logic tree. Responses are first categorized as work responses (the child attempts to do the task) or not-work responses. Each of these two responses is then divided into verbal and non-verbal categories. Within the work response category, both verbal and non-verbal responses are divided into two categories: (1) delimited responses, which do not go beyond the task, and (2) responses which are verbal or non-verbal extensions of work responses (spontaneous associations or other expressions in action or speech). Within the not-work response category, verbal not-work responses may be expressed in four ways: (1) competence (statements related to the child's ability to do the task), (2) negation (refusal to do the task), (3) aid (requests for help from the examiner), and (4) verbal substitution (verbal responses irrelevant to the task). In our use of the system, within the non-verbal not-work category, two types of responses are possible: (1) non-verbal substitution (motor responses which are irrelevant to the task) and (2) passive non-response.

Within the categories of responses described by the Birch procedure, it is possible to distinguish categories which are related to task orientation and categories which are related to preference for verbal and non-verbal responses. We focused upon these two broad areas of response types, which we thought might reflect changes in the children in the Ancona program. We were especially interested in the task orientation measures, as they provide some collaborative evidence relative to our behavioral ratings.

We distinguished the following six measures based on the Birch procedure categories which we considered to be related to task orientation:

(1) Percentage of work responses.

(2) Percentage of verbal not-work responses which were verbal substitutions (indicative of lack of task orientations).

(3) Percentage of verbal not-work responses which were competence statements.

(4) Percentage of verbal not-work responses which were requests for aid.

(5) Percentage of verbal not-work responses which were negations (indicative of lack of task orientation).

(6) Percentage of non-verbal not-work responses which were non-verbal substitutions. (The complement of this category is passive non-response, which is at least less clearly lacking in task orientation than is non-verbal substitutions.)

We distinguished four measures based on the Birch procedure categories which we considered to be indicative of preference for verbal resonse:

(1) Percentage of verbal items with work responses as compared to non-verbal items with work responses.

(2) Percentage of not-work responses which were verbal.

(3) Percentage of work responses with verbal extensions.

(4) Percentage of work responses with non-verbal extensions.

Our primary interest in examining these Birch procedure measures is to see whether the Headstart children show improvement in task orientation as a result of the Ancona experience. Secondarily we are interested in whether social-class differences are extant. Finally, we are interested in comparing our findings with those of Hertzig et al. and with our findings of last year for both methodological and substantive reasons. Methodologically, we have a reasonably good basis for comparing the Hawaii system (final responses only) with the Hertzig system to see if appreciable differences emerge. Substantively, we are interested in whether the social-class and ethnicity differences found by Hertzig (and confounded in their study) can be better understood. We will turn first to these last considerations and compare our findings with those of Hertzig et al. and our data from last year.

Comparison of the Ancona Results to the Results Reported by Hertzig et al.

Table 8 contains results from the Ancona 1968-69 research, results from the current year's research at Ancona, and the Hertzig et al. findings in relation to selected measures derived using the Birch procedure (mostly task orientation measures). The figures in Table 8 are computed for the children who were first year in school at the first testing, these children are most comparable to the working-class Puerto Rican and white middle-class children studied by Hertzig et al.

As is evident in Table 8, last year we found that the Ancona lower- and middle-class children were highly similar in their responses as categorized in this way. Furthermore, we found our sample to show a much higher level of work orientation than the Hertzig sample and to utilize Competence as the primary mode of not-work response. The findings on the sample last year were highly consistent with the results relative to the teachers' and testers' ratings which also indicated that the first-year children were highly similar regardless of social-class membership. In contrast, Hertzig _et_ _al_. found striking social-class differences in their sample, but these differences were confounded by ethnicity. Our data and their own analysis and discussion, suggested that the major differences found by Hertzig _et_ _al_. were more likely attributable to ethnic factors than to social class.

In comparing the responses of this year's first-year children at Ancona, we see some consistency with last year's data and some departures. In contrast to last year, there are some social-class differences. The middle-class children give proportionately more work responses than do the lower-class children. The middle-class children use Competence more often as a reason for not-working, and the lower-class children use Negation (refusing) more frequently than do the middle-class children. (The difference on negation just misses the usual levels for the statistical test.) Last the middle-class give significantly more verbal not-work responses.

The difference on negation is a reversal of their finding, the others are in the same direction of those found by Hertzig _et_ _al_. The Ancona children do not show a social-class difference in regard to Verbal Substitution whereas such differences were found by Hertzig _et_ _al_. Although more social-class differences emerge this year than did last year, the Ancona children are still strikingly more work-oriented than the Hertzig _et_ _al_. sample; this finding holds using the same procedure that was used in their study. In terms of not-work responses, the

Table 8

Comparison of Birch Procedure Results in the First Year Ancona Children (1968 and 1969) with Results Reported by Hertzig, Birch, Thomas and Mendez (1968)

	Ancona Samples						Hertzig et al. Sample		
	Lower Class	Lower Class	Middle Class	Middle Class	Signif. Levels for Social Class		Working Class	Middle Class	Signif Level Social Class & Ethnicity
	1968	1969	1968	1969	1968	1969			
Mean Age(Months)	45	39.1	45	44.8			42	40	
Task Orientation Measures*									
Proportion work responses	.898	.726	.890	.845	NS	.004	.64	.72	.001
Proportion verbal substitution	.246	.476	.180	.385	NS	NS	.64	.28	.001
Proportion competence	.740	.149	.781	.384	NS	.017	.10	.40	.001
Proportion aid	.014	.120	.005	.160	NS	NS	.11	.10	NS
Proportion negation	.000	.254	.034	.071	.10	.069	.15	.22	.001
Proportion non-verbal substitution	.436	.455	.303	.486	NS	NS	.43	.48	.05
Measure of Preference for Verbal Response									
Proportion verbal not-work	.555	.522	.760	.751	NS	.014	.38	.59	.001

---

*In the figures for the 1968 Ancona sample, the proportions are based on *final* responses only, whereas in the Hertzig *et al*. sample they are proportions of all responses. In the 1969 Ancona figures, all responses were also coded. However, in both cases for Ancona, the measures of proportion were computed for each child separately and then averaged. In the Hertzig *et al*. sample, the measures of proportion were computed for all responses of a group taken together.

samples are highly similar in regard to Competence and Aid and dissimilar in reference to Verbal Substitutions and Negations. The Ancona Headstart group uses Verbal Substitutions much less than the Puerto Rican working class group studied by Hertzig _et al_. This leads to a corroboration of their interpretation that the Verbal Substitution mechanism is primarily a product of ethnic group socialization. We do not know why the use of Negation is more prevalent in their middle-class group and less prevalent in our middle-class sample.

In comparing our findings last year with those of this year we must take account of both a change in method and the fact that we are dealing with a new sample. Assuming for the moment that methodology is the main factor accounting for differences, we do see some marked shifts. First of all, by using final responses only it seems clear that the proportion of work responses is somewhat overestimated: if a child is going to work on an item he will do so by the time you get his last response. Nevertheless, the Ancona children do tend to work more frequently than would be expected on the basis of the original Hertzig _et al_. study. The most striking alteration in the findings is in regard to the distribution of types of not-work responses. Last year, the overwhelming not-work response category was Competence, followed by some Verbal Substitutions and virtually no Aid or Negation responses. The noticeable decline in Competence responses in this year's data seems to be a function of methodology. Again, a final not-work response will tend to be a competence statement after other tactics have failed.

It should be recalled that we are dealing with a new sample this year, and that on other measures the first-year middle-class children and Headstart children appear less similar than did last year's group. Thus, some of the social-class differences found this year and the lack of same last year are likely to be a function of this change in sample. This is apparent in this year's data on the second- and third-year children, the bulk of whom were last year's first-year children. In comparing the older children on these measures, most scales show no social-class differences. The only difference is in regard to Verbal Substitution (the lower-class uses it more).

We believe the data we have collected using this procedure suggests that it is likely that social-class differences would appear on most of the measures if this technique were used on a wider scale. There do appear to be variations among the lower-class however, which deserve further study. Further, it seems clear that the ethnic confounding in the Hertzig _et al_. sample is contributing to certain differences, particularly the Verbal Substitutions measure.

Change in the Ancona Sample

Having compared our younger children with the Hertzig _et al_. sample, we turn now to the findings for the Ancona sample as a whole to examine any patterns of change which occurred from Test 1 to Test 2. Table 9 shows the ten scores of task orientation and preference for verbal response, computed for our lower-class and middle-class samples at each test time. We have already noted that the first-year children did show some differences in respect to task orientation, but their overall level of work-orientation was quite high. This finding seems quite consistent with our data on the ratings. The second-year children were more similar to one another than the first-year children, another confirmation of the rating information.

As can be seen in Table 9, when one combines the children into total lower- and middle-class groups, a number of social-class differences are evident on these measures in regard to the first testing. On the task orientation measures, proportion of work responses shows a significant difference, with the middle-class children giving a higher proportion of work responses. None of the other categories show differences significant at .05, but the differences on Verbal Substitutions (lower-class higher), Competence (middle class higher) and Aid (middle class higher) approach significance. In regard to the measures of preference for verbal responses, the middle-class children give significantly more work responses to verbal items, and more not-work responses which are verbal as opposed to non-verbal. In addition, the middle-class children's tendency to use more verbal extensions is almost significantly higher than that of the lower-class. These data seem to indicate some differences in our samples this year. The major area of difference seems to be in preference for verbal response and secondarily in regard to task orientation. As we noted earlier, the level of work orientation in our sample is quite high and differences should be considered in that light.

Having examined the initial status of the children, it is of interest to see if their behavior changes over the year on these measures. In the examination of ratings which tap similar factors, we found few changes in the children from the beginning to

Table 9

Birch Procedure Measures in Lower-Class and Middle-Class Groups at Two Test Times

	N = 29 Lower-Class		P Level	N = 29 Middle-Class		P Level	P Level LC/MC	P Level LC/MC
	Test 1	Test 2	Time 1-2 Diff.	Test 1	Test 2	Time 1-2 Diff.	Time 1 Diff.	Time 2 Diff.
Task Orientation Measures								
Percentage work	77.26	83.83	012	83.96	88.33	035	010	031
Percentage verbal substitution	43.56	47.55	NS	29.89	16.24	069	096	000
Percentage competence	32.55	38.50	NS	47.88	67.41	022	060	002
Percentage aid	8.4	7.18	NS	15.75	6.89	020	059	NS
Percentage negation	15.49	6.77	NS	6.48	9.46	NS	NS	NS
Percentage non-verbal substitution	36.61	36.53	NS	40.59	14.43	008	NS	017
Measures of Preference for Verbal Response								
Percentage work to verbal items	69.11	77.53		79.40	84.99	037	006	016
Percentage work to non-verbal items	87.99	93.67		92.25	95.55	053	NS	NS
Percentage verbal not-work	56.93	66.19		72.62	80.03	NS	034	011
Percentage verbal extensions	3.41	3.19		5.34	4.25	NS	091	NS
Percentage non-verbal extensions	1.12	.58		1.73	.75	033	NS	NS

the end of the school year. Last year we found few changes on the Birch procedure categories as well. Table 9 reveals that the Headstart children do change significantly in regard to work responses, becoming more work-oriented with the passing of a school year. On the other task orientation measures there are no significant changes. In regard to the measures reflecting preference for verbal responses, the lower-class children change significantly in the percentage of work to verbal items and in the percentage of work to non-verbal items. The proportion of work responses is greater to non-verbal items for our entire sample. We take these simultaneous changes to reflect greater responsiveness to both kinds of tasks.

The middle-class children show more change on the task orientation measures than did the Headstart group. As did the lower-class children, middle-class youngsters increase in their proportion of work responses. They also show a significant increase in the use of Competence as a reason for not working, a decrease in the use of Aid, and a marked decrease in the use of non-verbal substitutions. From data not presented in full here, it is clear that the increase in Competence is due to the first-year children as the second-year middle-class children are high on Competence to begin with (.614) and stay high. In the preference for verbal response area, the middle-class children significantly increase in the percentage of work responses to verbal items, their increase in percentage of work to non-verbal items is near significant, and their decrease on percentage of non-verbal extensions is significant (although the frequency is very low on both occasions).

At the end of the school year, more social-class differences are apparent in comparing total groups than was the case at Time 1. On the task orientation measures, the middle-class children give significantly more work responses, the lower-class make use of Verbal Substitution significantly more, the middle-class use Competence as a reason for not-working significantly more, and the lower-class children use proportionately more Non-Verbal Substitutions. These differences on the task orientation measures would all seem to favor the middle class. However, the proportion of work responses is very acceptable for the lower-class sample (84%) and this fact must be considered in interpreting the differential distribution of types of not-work responses. The lower-class children seem to prefer expressing not-work through the mechanism of Verbal Substitution (irrelevant verbalizations).and their distribution of types of not-work responses is quite consistent over the course of the year. The middle class change in their distribution of type of not-work responses, making increasing use of the Competence category with subsequent decreases in the use of Verbal Substitution and Aid. The sharp decrease in the use of non-verbal extensions on the part of the middle-class children seems to indicate their increasing tendency to deal with task demands through verbal means.

The pattern of differences at Time 2 on the preference for verbal response measures, is highly similar to that found at Time 1. The only social-class differences on these measures are in regard to the proportion of work responses to verbal items (the middle-class respond to more verbal items) and in the percentage of not-work responses which are verbal. This last difference is consistent with the decrease in the use of non-verbal extensions seen in the middle-class children.

Generally, the Birch procedure seems to reveal both more differences in the sample by social class, and more change in each group with time than did the ratings. There is some indication that the Headstart children do become more work-oriented as a function of their experience in school, but this is also true of the middle-class children and therefore their relative status remains virtually unchanged. Most of the changes revealed in the Birch procedure categories are desirable ones, indicating greater adaptation to the cognitive demand situation as reflected in the Stanford-Binet test. The middle-class children show one pattern of development which could not be tapped by the ratings and which may be significant in the long run. They seem increasingly to deal with work situations verbally, whether they are working or not-working they talk to the examiner and only occasionally communicate through actions alone. This is less the case for the Headstart children although they are also moving in this direction to some extent.

In regard to our hypotheses regarding school facilitative behaviors, there is some evidence of increases in desirable directions from both the ratings and Birch procedure analysis. Change as a function of the school experience seems more apparent in the Birch procedure analysis than it did in the ratings. It is still the case, as last year, that the level of work orientation and related skills seems quite acceptable in both groups even at the beginning of the school year. These generally high levels may account for the fact that change is not more apparent; we are almost dealing with ceiling effects to some extent.

The area which is still clearly problematic for the lower-class is verbal. Work responses to verbal items are less frequent than to non-verbal items and they continue to utilize non-verbal substitutions as a mechanism for not-working. These scales seem to indicate that the lower-class children although moving in this direction, have not yet accomplished the shift to the almost exclusive use of words in order to make their needs and desires known to an adult. They respond well to cognitive demands, working most of the time, but exhibit more actions as a way of not-working than do the middle-class children. Similarly, on the ratings, the area of Verbalization (understandability of speech) was the one showing greatest weakness in the Headstart group. It may be that the mild difficulty experienced by adults in understanding the children, leads the children to use the verbal mode somewhat less frequently than would otherwise be the case.

Social Interaction

Having discussed the cognitive and behavioral skills of the nursery school children, we turn now to our last major concern, the social interactions of the children. The Ancona School being integrated both in social class and race is a unique setting in which developing social interaction among these groups can be viewed. This year's study was a continuation of previous investigation about the following hypothesis: interaction between middle-class and disadvantaged children will increase with tenure in the program. Both cross-group social acts and friendship choices between groups will show increases in children with longer tenure.

An observational technique was used to determine both with whom the children were interacting and the gross quality of the interaction. Beginning in the fifth month of the school year, daily five minute observations were made of each sample child until a total of 100 minutes were collected. Following the observation, the social interaction recorded in these records was then coded on the basis of whether it was friendly, (dominative, integrative), aggressive (unprovoked, provoded, physical, verbal) or isolating behavior (ignoring or withdrawing behavior without evidence of aggressive intent. Only the first social interaction occurring in each minute was coded; thus the highest possible number of acts for each child was 100. Teacher-student interaction was also coded so that we might know the ratio of peer to teacher-student acts and the dependence upon the teacher shown by the child. The coding system is described in the methods section of this report. (See pp. 9-10.)

On the basis of a study done by McCandless and Marshall (1957) in which they found a moderate correlation between like choices on a sociometric rating and play choices in free play, it was decided to use observational data only and from this to compute best friend choices. A best friend was defined as that child with whom the observed child had the most friendly interactions. Because of this approach, information concerning dislike (negative) choices from the sociometric was not obtained. This information last year showed that among the second-year* middle-class children on the second test although there were more positive choices of lower-class Negro children than among first-year middle class, there was also an increase in negative choices of lower-class Negro children. And all middle-class children increased in negative choices of Negro children while among Headstart children this was not true. The reason for this contrast between positive and negative choices was not clear, but as negative choices of Negro children were not generalized to unknown Negro children, it seemed that the increase was reflecting both more positive and negative choices on an individualized basis. This year, in concentrating on obtaining an adequate record of the interaction occurring in the classroom, the total time of observation was increased to 100 minutes during which the observers attempted to record as much of the child's activity as possible. Thus for each child, these observations yield information as to the number and type of social interaction in which the child was involved and makes possible a relative comparison of these interactions with those of other children in the same social class and with children in another social-class grouping. The observations show with whom a child is interacting; thus some judgment about who are his best friends can be made.

---

*Throughout this report, second-year children designates those children with longer tenure and also includes third-year children.

Before looking at the results, it should be noted that the Headstart children in the sample live in the same neighborhood near the school and can often be seen playing together outside school hours. Also, the second-year children this year were among the first-year children of last year's sample.

An overview of the data shows that on the average about one act was coded every two minutes. And out of a possible total of 100 social interactions, the mean number of acts occurring for all children was 53.4. For middle-class children, the mean (59.0) was higher than that for Headstart (48.1). As might be expected, children with longer tenure in program (mean age in months, 55.1) had an average of 56.6 acts as compared to first-year children (mean age, 41.0) who had an average of 50.5 acts in 100 minutes. The predominant mode of interaction was friendly (87.6%). Aggression accounted for about 8.8% of the acts and isolating behavior, 3.6%. In ratios of friendly to aggressive acts, we found the Headstart ratio to be 9:1 as compared to 12:1 for middle-class children. Also, for Headstart there was a near equal distribution of physical and verbal aggressive acts while for middle class, 80% of the aggressive acts were verbal. Although not included in later discussion of data, we found that in terms of sex differences girls had a mean number of 51.5 acts compared with boys who had a mean number of 55.5 acts.

The teacher-student interaction, which was also coded on the basis of a possible one act per minute, showed that the overall average of acts was 17.6. Headstart children had an average of 19.3 acts and middle class, 15.9 acts out of a possible total of 100. McCandless (1957), found that adult dependency measured by number of social interactions directed towards adults in free play was negatively related to both sociometric status and observed social acceptance. We investigated as to whether children who had a low peer interaction also had a high teacher-student interaction. For Headstart children, of the 17 with low peer interactions (below the median), 14 of these children had high teacher-student interactions or 82%. Among the middle-class children, out of 9 with low peer interactions, 6 had high teacher-student interaction or 66%. Thus it seems that children who interact less with their peers also tend to have a high rate of interaction with their teachers; this relationship being most evident among the Headstart children.

Table 10 on page 58, shows the distribution of social acts and classroom distribution relative to each group. For example, the total Headstart group (n = 26) had in their surroundings 19.8% Headstart children, 35.1% middle-class Negro children and 45.0% middle-class white children to interact with. And the distribution of their total acts were 28.2% to Headstart, 39.4% to middle Negro and 32.0% to middle white. A Chi-Square test to determine if there was a significant difference between the distribution of social acts among the groups and their actual classroom distribution was computed and included in the table. For the example just given, this difference was significant at the .001 level.

These distribution breakdowns are shown for total Headstart and middle class and are further broken down by tenure in program (first-year vs second-year). For example, first-year Headstart children (n = 12) distribute 32.6% of their friendly acts to middle-class Negro children while first-year middle class (n = 14) distribute 47.1% of their friendly acts to the same group. The second-year middle class, in distribution of social acts, approaches most closely the actual classroom distribution of groups. All percentages are figured from the mean number of acts for each category over the total number of acts for all categories. Thus the percent of friendly acts of middle class to Headstart children is based on mean number of friendly acts to Headstart divided

Table 10

Headstart and Middle Class Percent Distribution of Social Acts

HEADSTART (26)

# acts	$\bar{X}$ acts	% Classroom distribution	HS (19.8)	MN (35.1)	MW (45.0)	Signif. of diff. from actual distrib. by X^2 test
1094	42.0	F:	24.5	34.3	27.3	
127	4.9	Agg:	3.1	3.5	3.2	
31	1.2	IB:	0.6	1.6	1.5	
1252	48.1	Total:	28.2	39.4	32.0	.001
FIRST YEAR (12)						
489	40.8	F:	27.0	32.6	25.5	
46	3.8	Agg:	2.3	2.8	3.9	
22	1.8	IB:	1.1	2.1	1.9	
557	46.4	Total:	30.4	37.5	31.3	.001
SECOND YEAR (14)						
605	43.2	F:	22.3	35.7	29.0	
81	5.8	Agg:	3.9	4.0	2.7	
9	0.6	IB:	0.2	1.2	1.1	
695	49.6	Total:	26.4	40.9	32.8	.001

MIDDLE CLASS (25)

# acts	$\bar{X}$ acts	% Classroom distribution	HS (24.3)	MN (32.8)	MW (43.8)	X^2 Signif.
1341	53.6	F:	15.6	41.5	31.9	
112	4.5	Agg:	2.3	3.5	1.9	
23	0.9	IB:	0.3	2.7	0.5	
1476	59.0	Total:	18.2	47.7	34.3	.001
FIRST YEAR (14)						
668	47.7	F:	11.9	47.1	26.2	
70	5.0	Agg:	3.0	3.7	2.6	
17	1.2	IB:	0.3	5.1	0.3	
755	53.9	Total:	15.2	55.9	29.1	.001
SECOND YEAR (11)						
673	61.2	F:	19.3	35.9	37.6	
42	3.8	Agg:	1.6	3.3	1.2	
6	0.5	IB:	-	0.3	0.7	
721	65.5	Total:	20.9	39.5	39.5	.01

Table 11

Headstart and Middle-Class Best Friend Choices

	HS	MN	MW	
All Headstart	42.3	34.6	23.0	.001
First year	50.0	33.3	16.7	.001
Second year	35.7	35.7	28.6	.001

	HS	MN	MW	
All middle class	28.0	44.0	28.0	.001
First year	28.6	50.0	21.4	NS
Second year	27.3	36.3	36.3	.01

HS: Headstart
MN: Middle-class Negro
MW: Middle-class white

by the total number of acts by all middle-class children.

Results for Headstart

Overall, as seen in Table 10, Headstart children differ significantly in the distribution of their social acts to Headstart, middle-class Negro and white children and the actual classroom distribution of these groups. A higher portion of their acts are with Headstart and middle-class Negroes than with middle-class white children. Acts to middle-class Negroes are close to the actual classroom distribution while those to Headstart are higher and those to middle-class white children are lower than actual distribution.

First- and second-year children do not differ significantly (.05) in the distribution of their acts between the groups. Second-year children have less acts to their own group with the difference going equally to middle-class Negro and white children. However, the percent aggression to their own group and to middle-class Negro children is greater in the second year while that to middle-class white children decreased.

Best friend choices of Headstart children are shown in Table 11, page 58. As described earlier, a best friend is that child with whom the observed child had the greatest number of friendly interactions. These results reflected the overall social interaction pattern. For Headstart children, the percent choice of best friends among middle-class Negroes was close to the actual classroom distribution (34.6% vs 35.1%). However, the percent best friends chosen from their own social class (Headstart) was considerably higher than the actual classroom distribution (42.3% vs 19.8%). Also, percent choice of best friends from middle-class white children was considerably lower than actual classroom distribution of this group, (23% vs 45%). In a comparison of Headstart first- and second-year children, children with longer tenure had a lower number of best friend choices among Headstart children (50.0% to 35.7%) and significantly higher percent best friend choices among middle-class white children (16.7% to 28.6%). Best friend choices among middle-class Negro children remained nearly the same. However, even with these changes, the percent of best friend choices among middle-class white children still was significantly less than their actual classroom distribution (28.6% vs 45%) and the percent best friend choices to Headstart was significantly higher than actual classroom distribution (35.7% vs 19.8%).

Summary

While the overall social interaction differs only slightly with tenure in program, there is a greater choice of best friends by the second-year Headstart children among middle-class white and fewer best friend choices from their own social-class group. This shift in best friend choice is accompanied by a change in distribution of aggressive acts. While second-year Headstart children had more percent friendly acts to middle-class white, they had less percent aggressive acts to middle-class white than the first-year Headstart children. Also, while the percent friendly acts of second-year Headstart to their own group was less than first year (22% vs 27%), the percent aggressive acts to Headstart was more (3.9% vs 2.3%). Thus although there was not a great shift in the number of total acts between these groups, there was a change in the quality of these acts. As best friend choices are based on friendly acts, these results are consistent with the expectation that as percent best friend choices from a group are higher with longer tenure, a lower percent aggression to this group should also be observed. This however is not true for the middle-class Negro group. Headstart second-year children had a slightly higher percent best friend choice of middle-class Negro than first-year children (35.7% vs 33.3%), but at the same time the percent aggression of second-year Headstart children to middle-class Negro was considerably higher than that of first-year Headstart to middle-class Negro (4.0% vs 2.8%).

Although the social interaction data on Headstart children did not strongly support the hypothesis that with longer tenure in program children will increase their cross social group acts, neither did it indicate that the opposite was true. However, on best friend choices children with longer tenure in program did show a greater proportion of cross social group choices than first-year children.

Relation to Last Year's Data

It is questionable as to how comparable the absolute values of last year's and this year's data are since different measures were used. However, an overall look shows that while in last year's data, second-year Headstart children had a significantly greater number of acts to their own group than first-year Headstart, this was not true this year. Last year the shift between first- and second-year percent acts directed to Headstart was 16.5% for the first year and 39.2% for the second year. In comparison this year's Headstart first-year children had 30.4% of their acts to other Headstart children while the second-year children directed 26.4% of their acts to their own social-class group.

Also best friend choices for thisyear do not reflect the differences found in last year's sociometric like choices. Last year, second-year Headstart children had significantly higher sociometric like choices within their own group than the first-year Headstart children. This year, however, second-year children had a lower number of best friend within group choices than the first-year children. In both this year and last year, the second-year Headstart children had almost the same percentage of best friend choices to their own social class (35.7% and 35.3%). However, there was a great difference between the first-year children of last year and this year. Last year, first-year Headstart children chose 15.9% of their like choices from their own social class while this year, first-year Headstart children chose 50.0% of their best friends from their own social class. In other words, although in both this year and last year the second-year children look nearly the same in terms of like/best friend choices, the shift to higher within-group like choices last year with longer tenure and the shift this year to lower within-group best friend choices with longer tenure are based on a large initial difference between last year's and this year's first-year Headstart children.

Results for Middle-Class

As shown in Table 10, in percent distribution of total acts to Headstart, middle-class Negro and white children, the middle class differs significantly from the actual classroom distribution with the greater proportion of total acts going to middle-class Negro children.

Significant differences are also evident between first- and second-year children. Those children with longer tenure in program have a higher percent of total acts to Headstart than first year (20.9 vs 15.2) and a higher percent total acts to middle-class white than first year (39.5 vs 29.6). This shift is also reflected in a significant decrease in percent total acts to middle-class Negroes; 39.5% as compared to 55.9% by first-year children to middle-class Negroes. Thus children with longer tenure have a greater portion of their acts to Headstart and middle-class white children than middle-class children new to the program. Distribution of total acts for second-year middle-class children approaches the actual percent classroom distribution of these groups.

Best friend choices, Table 11 for total middle-class children reflect the same story as the social interaction data. Middle-class children overchoose middle-class Negroes and underchoose middle-class white children as their best friends. Choice of best friends from Headstart and total middle-class children was near actual percent classroom distribution. Thus while the percent of total acts to Headstart was lower than the percent classroom distribution, the percent of best friends chosen from Headstart was slightly higher than percent classroom distribution (28.9% vs 24.3%).

In a comparison of first- and second-year children, the best friend scores also reflect the social interaction pattern. As in social interaction, children with longer tenure in program had a significantly lower number of best friend choices among middle-class Negroes than first-year children. Although best friend choices to middle-class white children increased as percent total acts did also, this was still less than the actual classroom distribution (36.3% vs 42.8%). Thus second-year middle-class children also were overchoosing Headstart and middle-class Negro children, but their percent best friend choices were nearer actual classroom distribution of these groups than first-year children.

In terms of the hypothesis, the middle-class children with longer tenure do have a higher percent distribution of cross social group acts than the first-year children in the program. Although a difference in best friend cross social group choices with tenure in program was not evident, the percent distribution of best friend choices among Headstart by both first- and second-year middle-class children was higher than the actual percent classroom distribution of Headstart children.

Relation to Last Year's Data

The results from this year and last year are similar. This year's second-year middle class had a slightly higher percent total acts to Headstart children than first year (20.9% vs 15.2%) which was also true the previous year. While this year's best friend choices do not show an increase of best friends across social-class lines with longer tenure as in the previous year, they do show a consistent percent choice of best friends from Headstart children near their actual distribution in the classroom.

Summary

Overall, middle-class children were interacting dispropotionately more with middle-class Negroes. However, with longer tenure in program there is a significant change in distribution of acts with greater percent total acts both to Headstart and to middle-class white children and fewer percent acts to middle-class Negroes. This was reflected also in best friend choices. In an attempt to better understand what was happening among the first-year middle-class children who were interacting with middle-class Negroes in a much higher proportion than their actual classroom distribution, middle class was further broken down by year and race. These figures are shown in Table 12.

It is clear that these first-year middle-class Negro children were highly active (mean number acts 65.4 as compared to 42.4 for first-year white children), and that they interacted mainly with other middle-class Negro children. It is interesting that first- and second-year white children also interacted disproportionately more with this group. When looking at the first-year middle-class Negro children individually, one finds that they are children with high total peer acts and that four out of seven are paired mutual friends. However, the explanation for the middle-class white children directing a disproportionate number of acts to these children seems to be different. Only 3 out of the 11 children have best friend choices within the middle-class Negro group, only 1 of which is a first-year child. Thus it seems that although the middle-class white children do not have a large percentage of their best friend choices from the middle-class Negro children, there is a great deal of social interaction between these groups within the classroom. A clearer investigation of these variances are, however, beyond the scope of this paper.

The group with which both middle-class and Headstart children

Table 12

Negro and White Middle Class Percent Distribution of Social Acts

NEGRO (14)

			HS	MN	MW	Signif. of diff. from actual distrib. by X^2 test
% Classroom distribution			(24.3)	(30.6)	(45.0)	
# acts	$\overline{X}$ acts					
826	59.0	F:	16.1	41.6	32.0	
76	5.4	Agg:	2.1	4.4	2.1	
14	1.0	IB:	0.3	0.9	0.6	
916	65.4	Total:	18.5	46.9	34.7	.001
FIRST YEAR (7)						
406	58.0	F:	11.9	52.4	24.1	
44	6.3	Agg:	2.0	5.2	2.3	
8	1.1	IB:	0.6	1.2	0.2	
458	65.4	Total:	14.5	58.8	26.6	.001
SECOND YEAR						
420	60.0	F:	20.3	30.7	39.9	
32	4.6	Agg:	2.1	3.5	1.9	
6	0.8	IB:	-	0.5	1.1	
458	65.4	Total:	22.4	34.7	42.9	.10

WHITE (11)

			HS	MN	MW	X^2 Signif.
% Classroom distribution			(24.3)	(35.1)	(40.5)	
# acts	$\overline{X}$ acts					
515	46.8	F:	15.0	41.4	31.7	
36	3.3	Agg:	2.5	2.3	1.5	
9	0.8	IB:	0.2	5.1	0.4	
560	50.9	Total:	17.7	48.8	33.6	.001
FIRST YEAR (7)						
262	37.4	F:	13.4	39.3	30.7	
26	3.7	Agg:	3.5	2.0	2.4	
9	1.3	IB:	0.2	7.9	0.6	
297	42.4	Total:	17.1	49.2	33.7	.001
SECOND YEAR						
253	63.2	F:	17.6	45.0	33.6	
10	2.5	Agg:	0.8	3.0	-	
-	-	IB:	-	-	-	
263	65.7	Total:	18.4	48.0	33.6	.001

have the least interaction is that of the middle-class white children. Only the second-year middle-class Negro children in their social acts approach the actual classroom distribution of the middle-class white children. A partial explanation is reflected in the fact that these children have the lowest mean number of social acts. Goodman (1952), also found that in a biracial nursery school, white children had a lower involvement in social interaction than expected by chance. This was thought to be due to differences shown in personality ratings in which Negro children were more gregarious, vigorous, aggressive while white children were more socially and physically apprehensive and rated higher on emotional control, conformity, patience and affectionateness. No statement concerning personality differences can be made concerning the Ancona children.

In conclusion, the data from this year's research shows that although there are instances of disproportionate distribution of social acts among the three groups discussed, the social interaction among the children at Ancona is for the most part distributed across class and racial groups. Children do not tend to be isolated into groups by race or social class. The setting at Ancona appears to be an open and facilitating atmosphere in which children, able to move about, working individually or in groups have much opportunity for a variety of social encounters. It is suggested that to better understand the significance of the distribution of these acts, it would be necessary to investigate in greater depth the individual children involved. From this, one might then learn more about the basis of individual friendship, dislike and neutral choices and the role of these interactions in the development of the children involved.

PART II

THE STUDY OF ELEMENTARY SCHOOL CHILDREN AND SIBLINGS WITHIN FAMILIES WITH CONTINUED ENROLLMENT

This section is concerned with the effectiveness of the Ancona Montessori program on the educational achievement of children who continue their schooling into the primary grades at Ancona. We are also interested in the possibility of diffusion effects between siblings within families due to continual enrollment of children from the same families in the program. We have maintained the hypotheses of previous reporting periods which reflect our optimism that length of participation in the program will have positive effects for the individual child, and that in continuing families of Ancona children, a diffusion effect will operate to increase cognitive and school readiness skills of older and younger siblings.

Hypotheses to be Investigated

The following hypotheses have been previously investigated and discussed in the 1968-69 final report (Stodolsky and Jensen, 1969). This previous study suggested that hypotheses concerning the effectiveness of continued enrollment can be tentatively upheld. Diffusion effects to siblings of continuing families were not much in evidence for a limited sample of younger siblings entering the nursery program. The younger siblings did have a slightly higher initial I.Q. and seemed somewhat more inhibited than non-siblings. From a sample of Ancona children who were attending the public primary school at that time, those children who had younger siblings attending Ancona were performing better than those who had no siblings attending Ancona. Similar hypotheses are under investigation this year; they are as follows:

Hypothesis 1c: Disadvantaged children who continue in the Montessori elementary program will show higher school achievement than those who go to public school.

Hypothesis 3: Continuing involvement of the same families in the program will result in "diffusion effects" to the intellectual attainments of older and younger siblings.

Hypothesis 3a: In examining children who attended Ancona at one time, older siblings will show school attainment that is better than non-siblings in public schools.

Hypothesis 3b: Younger siblings entering the program will show higher initial cognitive and behavioral skills.

Hypothesis 1c: The Sample and Measures

The sample for this investigation was obtained by tracing public school records of children previously at Ancona and now attending the public schools, and by testing the primary school children remaining in the Ancona program.

A. Children now attending public school. This sample can be divided into two groups:

Group I

Fifteen of the children who attended the 1965 summer eight week program and then went on to public school are presently in the fourth grade. Of the 15 children located last year who spent eight weeks at Ancona only three children who were put into special ungraded status have additional test data. One child is deceased, one parent refused to release a child's public school records, and four were not traceable. Because children are tested in the third grade of public school it is possible to assume that the four children who were untraceable this year have no additional

Group II

This group contains ten children who have spent from one to two years in the Ancona nursery program and now are attending public schools. Members of this group are currently in grades K through third. Public school records for this group yielded Metropolitan achievement data for the three children now in the third grade who had all spent two years in the Ancona nursery program. On three different children who also spent two years at Ancona and are now in the first grade we have Kuhlman-Anderson I.Q. scores plus Metropolitan Reading Readiness test scores. The two kindergarten children received no standardized tests, and on two children who spent one year at Ancona and are now in the first and second grade, no new tests were administered.

For comparative purposes then, this sample is reduced to six children, all of whom spent two years at Ancona; three are in the third grade and were tested with the Metropolitan Achievement battery and three are in the first grade and were given the Metropolitan Readiness and the Kuhlman-Anderson I.Q. tests.

B. Children now attending Ancona primary classes.

Ten children are continuers in the Ancona elementary program. The classrooms are ungraded but grade estimation can be based on the age of the child. Two of the children have been in the Ancona program for five years, four for four years and four for three years. These children were administered the Metropolitan Achievement battery in two forms near the end of the school year. Six received the battery for grade one and four received the grade two battery. At approximately the same time, all ten children were administered five sub scales of the Wechsler Intelligence Scale for Children: (1) Block Design, (2) Arithmetic, (3) Object

Assembly, (4) Coding, and (5) Mazes. These ten children were also rated on non-cognitive variables on the project's teacher rating scale of behavior at the beginning and end of the year.

Hypothesis 1c: Results and Discussion

The effects of length of enrollment at Ancona will be investigated by reporting information available from the variety of instruments mentioned on the groups or subsamples just described. Comparison of performance on the instruments will be discussed where it is possible.

Group I: (Children who were enrolled in the 1965 eight week summer program and now attend public school.)

Previously reported Metropolitan Achievement scores for this group of children has found them to be below grade level on all subscales. A subsample of children who have siblings continuing in the Ancona program did better than non-siblings. The only new data for this group comes from the public school records of three children who were put into special ungraded status by the public school for remedial purposes. These children were given the Metropolitan Achievement Battery at the end of the school year. Their scores are shown in Table 13.

Table 13

Metropolitan Achievement Grade Equivalent Scores for Three Children Who Were Placed in Ungraded Public School Classrooms

	Month	Age	Word Knldge.	Word Disc.	Read.	Spell.	Lang.	Ar. Comp.	Ar. Prob.
Child A	4/70	9.7	4.7	4.0	4.0	5.5	4.9	4.3	3.7
Child B	4/70	9.4	1.6	2.5	1.6	1.8	1.4	2.3	2.3
Child C	4/70	9.9	2.8	3.6	2.2	3.8	3.5	3.3	2.3
Means			3.03	3.36	2.6	3.7	3.26	3.3	2.7

The mean scores for the entire Group I's Metropolitan Achievement data when the children were tested in the third grade (1968-69 school year) are shown in Table 14. By subtracting the original scores of the three children just discussed and substituting their retest scores (1969-70), a new mean can be calculated for the fifteen children. These up-to-date figures are also shown in Table 14.

Table 14

Metropolitan Achievement Grade Equivalent Scores for Children in Third Grade

	Word Knldge.	Word Disc.	Read.	Spell.	Lang.	Ar. Comp.	Ar. Prob.
1968-69 Testing (N= 15)	2.2	2.4	2.2	2.7	2.4	3.0	2.5
1968-70 Testing (N= 15)	2.4	2.7	2.5	2.9	2.6	3.1	2.7

Children in Group I perform below grade level on all scales. By adding improved retest scores for three of the originally lowest scoring children, the group's mean scores improve but remain below grade level. As a group these children are about one year below expected grade level. This decrement is rather typical of disadvantaged elementary school children (Bloom, Davis and Hess, 1965). It seems clear that the eight-week intervention at the preschool level does not have long-term effects on the children's performance.

Group II: (Children who have spent one to two years in the Ancona nursery program and are now in public elementary school.)

Three third grade children in this group were administered the Metropolitan Achievement battery. All three had been at Ancona for two years. Their scores are shown in Table 15.

Table 15

Metropolitan Grade Equivalent Scores For Three Ancona Graduates Now in Third Grade

	Age	Word Knldge.	Word Disc.	Read.	Spell.	Lang.	Ar. Comp.	Ar. Prob.
Child A	8.4	1.1	2.4	1.3	1.8	3.4	3.4	-
Child B	8.6	2.7	2.6	3.1	3.5	2.2	3.6	-
Child C	8.10	2.3	2.2	2.4	2.6	3.1	3.5	2.3
Means		2.0	2.4	2.3	2.6	2.9	3.5	-

The scores for this limited number of children who spent two years at Ancona, show that the children are generally performing below grade level except in Arithmetic Computation. Their mean performance on the scales closely resembles the scores of children in Group I who had only eight weeks of experience at Ancona. Longer preschool tenure does not seem to influence the later performance of the children, with the possible exception of arithmetic computation. Number skills are stressed in the Montessori preschool curriculum.

Three first-grade children were administered the Kuhlman-Anderson Intelligence Test and the Metropolitan Reading Readiness Test. These children had all previously been in the Ancona nursery program for two years. Their scores are shown in Table 16. In addition Table 16 contains the mean performance for six children who were in first grade in 1968-69 and had spent from one to two years in the Ancona nursery program.

Table 16

Kuhlman-Anderson I.Q. and Metropolitan Reading Readiness Scores of First Graders

	Age	M.A.	I.Q.	Readiness %tile
Child A	7-1	7-3	102	48
Child B	6-3	6-5	103	19
Child C	6-2	6-11	112	55
Mean (1969-70) N= 3			105.6	40.7
Mean (1968-69) N= 6			103.2	29.0

The performance of this year's group is slightly better than last year's, but the differences are not dramatic and the numbers are too small to attempt further interpretation. Two of the three children in the first grade appear to be fully ready for school work at the elementary level.

Group III: (Head Start Continuers in the Ancona program.)

The effects of length of tenure on school achievement for this group will be investigated by presenting data for the instruments given to the child and attempting to make further comparisons.

The Metropolitan Achievement Data

The rationale behind giving this instrument to Group III children was based on an attempt to compare their age-grade status with those children from the other groups now attending the public schools. The major difficulty is that public school children are administered the instrument in the third grade. Furthermore, the non-graded classrooms of the primary grades at Ancona do not assign grade status to children. The director of the school administered the test and reported that children had difficulty in following group instructions. She attributed this to their lack of school experience in this area, from the Montessori emphases on self-directed activities and individualized instruction. She also felt the tests themselves were seen as threatening by the children. Three children refused to do one or more of the tasks demanded of them.

The results from this testing are reported in Table 17.

Table 17

Metropolitan Achievement Grade Equivalent Scores For Children Continuing in the Ancona Primary Grades N= 10

Battery I

	Age	Years in Program	Word Knowledge	Word Discrimination	Reading	Arithmetic Concepts and Skills
Child A	6-7	3	1.1	1.2	1.5	1.4
Child B	7-5	4	1.7	1.5	1.7	1.4
Child C	7-1	4	1.6	refused	1.1	-1.0
Child D	7-2	3	1.9	2.3	1.5	1.6
Child E	7-5	3	1.3	1.3	1.2	1.3
Child F	6-7	3	1.3	1.3	1.2	1.3

Battery II

	Age	Years in Program	Word Knowledge	Word Discrimination	Reading	Spelling	Arithmetic Concepts and Problem Solving and Computation
Child G	8-4	5	1.9	2.2	1.7	-1.0	2.7
Child H	8-4	5	1.4	1.4	-1.0	refused	1.1
Child I	7-11	4	refused	refused	refused	1.6	2.4
Child J	7-4	4	1.7	1.7	1.7	-1.0	2.0

As is apparent from Table 17, the children in this group are generally performing below expected grade levels. In examining the available test data last year for three children then in the Ancona first grade, we were cautiously optimistic about the prognosis for continuing children. This year's data on a larger group of children does not support our optimism. Most of the children are not making satisfactory progress in the Ancona elementary program.

It may be that for some reason the test administration was seen so negatively by the children that their scores are depressed, but this in itself would be symptomatic of school adjustment problems. We do know that Child I was experiencing a general upset at the time of testing and would expect her performance to be quite adequate on another occasion. This is corroborated by the fact that she performed extremely well on the WISC (I.Q. = 129). Child H, although eight years old, has only been placed in the elementary program for the first time this year. He is a seriously disturbed boy and has made a good adjustment to the elementary classroom.

It is probably more likely that special efforts will have to be employed with these children if achievement in these areas is desired. It is the case that some of the skills measured in the Metropolitan are not emphasized in the early part of the elementary program at Ancona. We turn now to the data on the WISC which we believed to be somewhat closer to the objectives emphasized.

The rationale behind administering the WISC subscales was the attempt to make an evaluation of performance on measures which would more closely articulate to the Montessori curriculum. Thus Arithmetic was administered because of the amount of time the children seemed to spend doing Montessori mathematics activities, and Block Design and Object Assembly seemed to tap a wide range of skills based on visual-motor integration which is reflected in the Montessori sorting, matching, and puzzle activities. Mazes was given to test the effectiveness of the Montessori training in eye-hand coordination exercises with pencils, for example inset drawing and map tracing. Coding was added in order to have enough scale scores to compute a prorated Performance I.Q. and because it also seemed to utilize eye-hand coordination skills.

Table 18 contains the ten children's prorated Performance I.Q. and test age equivalents of their subscale scores.

Table 18

Test Age Equivalents For WISC Raw Scores and WISC Performance I.Q. (prorated) On Ten Primary Head Start School Children at Ancona 1969-70

	Sex	Years in Program	Age	Performance I.Q.	Arith-metic	Block Design	Object Assembly	Coding	Mazes
Child A	F	3	6-7	111	7.2	7.6	8.10	9.2	4.10
Child B	M	4	7-5	113	6.2	7.6	13.10	6.6	7.6
Child C	F	4	7-1	85	5.2	5.6	6.10	6.2	4.10
Child D	M	3	7-2	94	8.6	5.6	6.10	6.6	8.2
Child E	F	3	7-5	93	6.2	6.6	4.10	9.6	6.6
Child F	F	3	6-7	100	6.2	7.6	6.2	6.6	6.6
Child G	M	5	8-4	90	7.10	10.10	8.10	6.2	7.6
Child H	M	5	8-4	71	5.2	6.6	4.10	6.6	6.2
Child I	F	4	7-11	129	8.6	12.10	10.10	11.2	9.10
Child J	M	4	7-4	114	7.2	8.10	7.10	11.2	7.10
Means			7-5	100	6.10	7-11	8-0	7-11	7-0

The performance of the ten disadvantaged children in the Ancona elementary program on the WISC does present a somewhat brighter picture than the Metropolitan data. The prorated Performance I.Q. for the group is normal and all but two children (Child C and Child I) test in the average range. Interestingly, these two children are sister and brother. The performance of the group on the subscales lends some support to our contention that these scales would more accurately reflect the emphases in the curriculum. The children perform above the standards for their age on Block Design and Object Assembly (visual-motor integration) and on Coding (eye-hand coordination). They are somewhat below age norms on Mazes and Arithmetic. We are particularly puzzled by the arithmetic performance, but believe the performance of the children reflects the fact that most have basic mastery of numbers but have not dealt with "word" problems which enter the scale in this age range.

Teacher Ratings of Non-Cognitive Variables Early and Late in the School Year

The project's teacher rating form was used to assess some aspects of the children's behavior early and late in the program. Pre- and post-test ratings are shown in Table 19..

Table 19

Teachers' Ratings of Behavior of Ancona Elementary School Children Early and Late in the 1969-70 School Year

Pre-Teachers' Ratings	Distractibility	Activity Level	Initiative with Materials	Self-Confidence on Tasks	Persistence	Reaction to Failure	Sense of Intellectual Challenge	Fear of Adult	Social Initiative with Adult	Communication of Affect	Compliance with Adult	Verbalization
Child A	6	5	7	5	4	5	5	5	5	3	2	6
Child B	8	3	5	8	8	1	6	1	3	2	3	5
Child C	6	4	4	5	6	5	5	6	4	5	6	3
Child D	6	5	4	6	5	5	6	4	3	3	6	5
Child E	8	2	4	8	8	7	8	7	3	1	8	6
Child F	6	6	3	6	5	5	6	7	4	6	5	5
Child G	7	2	4	4	5	5	4	2	5	4	6	1
Child H	9	2	9	9	8	9	8	4	3	1	8	6
Child I	3	3	3	6	4	4	3	7	7	6	2	5
Child J	5	3	4	5	7	9	5	3	4	4	4	3
Means	6.4	3.5	4.7	6.2	6.0	5.5	5.6	4.6	4.1	3.5	5.0	4.5
Post-Teachers' Ratings												
Child A	6	5	7	4	6	5	6	7	6	6	2	7
Child B	8	3	8	7	7	5	6	4	3	3	7	6
Child C	6	5	3	6	4	4	6	5	3	5	4	5
Child D	3	5	4	5	4	4	3	5	4	5	4	5
Child E	6	3	6	6	4	6	6	4	4	3	6	6
Child F	5	6	5	5	5	5	4	2	3	3	5	6
Child G	4	3	4	5	4	4	2	4	4	4	5	3
Child H	8	3	6	8	6	7	7	3	3	2	6	5
Child I	4	5	3	4	4	4	6	7	6	6	5	6
Child J	5	5	6	6	6	4	5	6	4	4	3	3
Means	5.5	4.3	5.2	5.6	5.0	4.8	5.1	4.7	4.0	4.1	4.7	5.2

Overall the ratings place the children within normal limits on these scales. Mean differences in scale points from pre- and post-ratings do not exceed one scale point with the exception of Persistence which moves in the direction of more persistence on hard tasks. On the Attention-Distractability scale the difference is .9 in the direction of less distractability. Other changes, such as in Activity Level are also in the desired direction.

An attempt has been made to investigate the hypothesis that children who go on to elementary school in the Ancona program will have higher school achievement than do those who go into public school. From examining the available data, the WISC findings seem favorable. On the performance subscales of the WISC, Block Design, Object Assembly, and Coding, the children as a group score four and five months above their expected performance for a group of their mean age.

On the negative side is the low performance of the children on the Metropolitan Achievement Battery. The other serious issue raised in the Metropolitan testing was the tester's report of the difficulty that the children had in following group given instructions and some children's refusal to participate on some of the subtests.

A critical question concerning the WISC data is the issue of comparability to children in public schools, and the failure of the group to achieve their mean age equivalent on the Arithmetic scales.

Although we have attempted to examine all data available to us, small numbers and lack of comparability have frustrated our efforts to some extent. We cannot be sure yet what the long-term picture for these children will be, but we do not believe much optimism is called for unless further efforts are made on their behalf.

Investigation of Diffusion Effects to the Public School Achievement of Older Siblings (Hypothesis 3a)

The current data add only a footnote to the 1968-69 study, in which it was found that a subsample (N = 5) of older siblings with younger siblings continuing in the Ancona program did better on the Metropolitan Achievement battery than did a subsample (N = 10) of children with no siblings. Because our data on Metropolitan Achievement is dependent on public school administration of the test which usually occurs in the third and sixth grade, this year two new siblings have reached the third grade where they received the Metropolitan Achievement Battery along with one non-sibling. In the previous year's sample all fifteen children had been participants in the 1965 eight week program. The three children reaching third grade this year have all been at Ancona for two years before their entry into public school. Their Metropolitan Achievement scores are shown in Table 15.

Child A, the non-sibling does considerably below grade level in word knowledge, word discrimination, reading, and spelling while the siblings B and C do better. To some extent these new cases tend to confirm the hypothesis that older siblings of continuing families who attend the public school do better than non-siblings in the public schools. However, it is well to note that for all of the Ancona children in public schools, the Metropolitan scores are usually below the grade norms.

Study of Diffusion Effects to the Cognitive and Behavioral Skills of Younger Siblings (Hypothesis 3b)

Among the thirteen first-year Headstart children in the 1969-70 sample, seven were younger siblings of children who were in the program or who had been at Ancona at one time. Of these seven, two children had siblings who are now in the public primary school and six had siblings in the primary and nursery classes at Ancona. One child has two siblings, one in public school and one at Ancona. Table 20 contains the entering 1969 Stanford-Binet data for this group. Contrary to the hypothesis of a diffusion effect to the intellectual attainments of younger siblings, the non-sibling group has a higher mean I.Q. score. This year's sibling group is comparable to last year's (mean 86.9) but the non-siblings are higher than last year's non-siblings.

Table 20

Binet Scores At Entrance Into the Ancona Program
For Younger Siblings and Non-Siblings in September 1969-70

Siblings N= 7

	C.A.	M.A.	I.Q.
Child A	3.7	3.3	89
Child B	3.2	3.4	102
Child C	3.2	2.3	71
Child D	3.0	2.9	89
Child E	2.7	1.11	74
Child F	2.8	2.2	81
Child G	4.6	4.4	95
Means	3.31	2.80	85.85

Non Siblings N= 6

	C.A.	M.A.	I.Q.
Child A	2.1	2.7	89
Child B	3.1	3.6	90
Child C	3.2	2.11	90
Child D	3.8	3.8	98
Child E	4.2	4.3	101
Child F	3.5	2.7	75
Means	3.31	3.20	90.50

In keeping with our attempt to make evaluation more appropriate with the Montessori Curriculum, children under four years of age were tested with the Merrill-Palmer Scale of Mental Tests. The groups of siblings and non-siblings can be compared on this measure also. However, two children of both groups were above the age level for receiving the test and one child refused the test and after several attempts were made to test him with this instrument, the child was dropped from this sample: Table 21 reports these data.

Table 21

Merrill-Palmer Mental Age, I.Q. and Percentile Rank of Sibling and Non-Sibling Entrants to the 1969-70 Ancona Program

Siblings N= 5

	M.A.	I.Q.	Percentile Rank
Child A	36	109	80
Child B	40	105	70
Child C	37	97	45
Child D	38	103	60
Child F	47	109	75
	39.6	104.6	66.00

Non-Siblings N= 5

	M.A.	I.Q.	Percentile Rank
Child A	41	121	95
Child B	47	102	60
Child C	38	100	50
Child D	46	102	60
Child F	40	98	45
	42.4	104.6	62.00

Data from the Merrill-Palmer Scale does not support the hypothesis of a "sibling diffusion effect" for younger children of continuing families. The groups are virtually identical in performance.

To assess the possible diffusion effects on younger sibling's non-cognitive behavior, means on the scales of examiner's ratings were compared to the non-sibling's means on examiner's ratings, which were made in conjunction with the Stanford-Binet given to both groups at entrance into the 1969-70 program. Table 22 contains these data.

Table 22

Mean Scores on Stanford-Binet Examiner's Ratings For Siblings and Non Siblings Entering the 1969-70 Program

	Distractability	Activity Level	Speed of Response to Verbal Items	Speed of Response to Performance Items	Initiative with Materials	Self Confidence	Persistance	Reaction to Failure	Sense of Intellectual Challenge	Willingness to Continue	Fear of Adults	Social Interaction With Adults	Communication of Affect	Compliance With Adults	Verbalization
Siblings N= 7	5.28	3.70	5.28	3.70	3.57	5.57	5.42	2.00	5.42	5.00	4.28	4.42	3.57	5.00	6.85
Non Siblings N= 6	6.66	3.33	5.33	4.0	4.0	4.0	5.33	1.66	6.83	5.83	3.50	3.00	2.66	5.50	5.66

Only four scales show differences of one point or more: (1) The non-siblings are more distractible by a difference of 1.38 which moves them slightly away from normal attentiveness (scale point 5) toward "attracted to things external to the task, but able to return attention." (2) The siblings are less self-confident by one-and-one-half scale points but still well within the normal range on the scale, while the non-siblings are only one scale point away from being rated "quite confident." (3) Siblings seem to respond with more sense of challenge, while the non-siblings' mean almost rated "apparently more comfortable with easy tasks." (4)In adequacy of verbalization non-siblings are closer to more easily understood and adequate speech whereas siblings' speech is rated more difficult to understand. On all other ratings mean difference between the two groups is usually half a scale point or less. There is no systematic evidence in these data that suggests supporting the hypothesis that non-cognitive variables as measured by our rating scales are affected by a sibling diffusion effect for younger members of continuing families. The differences wehave found, with the exception of Verbalization, might be interpreted as reflecting higher family expectations. Less distractibility, more intellectual challenge and concomittant loss of self-confidence might be a result of achievement pressure.

Conclusion

On the basis of these data, we do not believe a diffusion effect is operating in regard to younger siblings' of continuing families at Ancona. Birth order effects are an obvious confounding factor in assessing these data.

OVERVIEW

The research on Headstart children at Ancona began in 1965, a period of great hope in regard to the efficacy of preschool intervention for improving the life chances of disadvantaged children. Government support of the work is terminating with this writing, five years later. The current mood in the field is one of much more caution and recognition of complexities still only barely understood. This small research project has mirrored many of the trends seen more broadly.

It is nevertheless true that the individuals involved at the school, the teachers, parents and children remain committed to the idea of an integrated school and have decided to continue to enroll Headstart children in the program at the risk of considerable financial strain on a small private institution. This decision reveals what the "hard" data misses: the personal and community value placed on the program over the last five years.

Although it is possible to draw some relatively firm conclusions from this research, in a real sense the value of the program remains unknown to the researcher. On one level, without a "control" group of similar children, the effects on the Headstart children remain clouded. More idiosyncratically, the effects on the individual children had they not attended the program are simply unknowable.

From a researcher's perspective, the "hard" data do tend to confirm certain propositions and we will attempt to summarize these briefly.

1. Cognitive change. Attendance at the Montessori preschool does result in initial increases in intelligence test performance

as measured by the Stanford-Binet. Subsequent increases with continued years in the program are not found. For a number of years, we investigated the effects of participation on a number of Piaget tasks. We are convinced that the program does not affect the developmental course as reflected in those tasks. This year we attempted to come closer to an assessment of certain "performance" aspects of cognitive functioning which we believe more adequately reflect the emphases in the Montessori curriculum. The data on WPPSI, Merrill-Palmer, and WISC scales for the Ancona sample support this idea. Change on these tests (or above average performance for the elementary children) is present in children not only after their first school experience, but in subsequent years. This point needs further investigation, but we tentatively conclude that the Montessori curriculum is effective in such areas as the development of visual-motor integration capacities, sorting and matching skills, eye-hand coordination, and to some degree number concepts.

The data on the Stanford-Binet and other intelligence tests taken together points to both the strengths and weaknesses of the curriculum. The strengths have just been enumerated. The weaknesses seem to lie primarily in the verbal area. The data on intelligence tests, the rating scales, and the Birch procedure categories all indicate some verbal deficit on the part of the children, a deficit which is not appreciably reduced by attendance at a Montessori school.

2. School facilitative behaviors. This area has been studied by the use of ratings and other procedures, including the Hertzig-

Birch category system. The first groups attending the Headstart programs at the school were rather extreme on some of these scales, notably Distractibility, and showed marked improvement which correlated with improvement on intelligence tests. Groups studied in the last two years seemed more ready for school on these scales and therefore few social-class differences were found. The relatively satisfactory status of the children to begin with also resulted in little change from the beginning to the end of the school year. We cannot really explain the seeming shift in the composition of the sample over the course of this project.

3. Elementary school progress. Only a small number of children have to date continued on into the elementary program at Ancona having completed the preschool experience. We believe it is really too early to judge the long-term progress of these children, especially relative to their cohort in the public schools. It is nevertheless the case, that the data we have collected do not portend well for these children, in conventional terms. The children do not perform at grade levels on the Metropolitan Achievement battery, indicating that they are not mastering the usual first-grade learning at the rate which is expected nationally. Whether the timing of the elementary program is such that spurts will occur later is not known. Data on the same children on the WISC shows normal to above average performance in regard to some "performance" areas of functioning. These data taken together suggest that the elementary children do have areas of strength, but the elementary curriculum may not be building upon them. An

effort in this direction of curriculum building would seem appropriate. It may be that if the school wishes to accomplish the usual school learning (as reflected in the Metropolitan) they may be able to do so by revising their instruction to move from the strengths of the children or they might consider a more heavily linguistic program which is essentially compensatory.

4. Diffusion effects. Children who attended Ancona at one time and have younger siblings still in the School are achieving in the public schools at a somewhat better level than non-siblings, but neither group is at grade level. Again, the numbers of children involved here are small, but the data are consistent. No such diffusion effect seems evident in regard to younger siblings. They do not look more ready for school or brighter when they enter school than do non-siblings (if anything the reverse may be indicated). We cannot readily explain this differential diffusion effect. Birth order is clearly a confounding factor, but it may be that for older children who have already been through the program, the continued involvement is somewhat reinforcing, whereas the younger children not having yet had a school experience remain essentially unaffected.

5. Social interaction. Observations of children's interactions at Ancona have typically been characterized by a high level of interaction across social class and racial lines. The integrated setting does in fact lead to contacts among children regardless of their home background and racial identification. Data over the years suggests that there is some tendency for middle-class children to increase their interactions across racial

and social class lines with longer tenure in the program; whereas lower-class children may tend to make more contacts both positive and negative with members of their own group as time passes. On the basis of sociometric data we feel quite confident that the children are not shifting in their choices of friends or interaction patterns on a stereotypic basis, but rather that particular personality traits of children lead to these shifts. Such correlates have not been studied by us, but would seem a useful area for further investigation.

Recommendations

1. Future research. The differential effects of the Montessori curriculum on areas of cognitive functioning deserves further attention. Better, and more extensive, assessments of the performance and linguistic areas seems called for, as does an inquiry into the work styles developed by children under such a regimen. Experimentation with elementary school programs which articulate well with the children's strengths and weaknesses at the termination of a Montessori preschool program is badly needed. It raises the question as to whether a heavy performance emphasis can be used to teach traditional early academic skills.

More research is needed in the area of social attitudes and interaction among children of diverse backgrounds. Specific factors which lead to continuing positive interaction across groups should be isolated, and the role of the school and teachers in this process should be investigated.

2. Efforts specifically at Ancona. We believe the time is ripe for the staff to turn to a curriculum development project

based on what we have learned and a consideration of the educational objectives which they believe to be worthy. Greater effort can certainly be placed on attempts at individualizing instruction in the elementary program on the basis of differential readiness of the children.

The decision of the school to maintain the Headstart program is a sound one both educationally and socially. It is an opportunity to continue a positive relationship with the surrounding community, to give the children from all groups needed exposure to one another, and an intellectual challenge to the school and its staff.

BIBLIOGRAPHY

Bereiter, C. and Engelmann, S. Teaching Disadvantaged Children in the Presschool, Prentice-Hall, 1966.

Bereiter, C. Acceleration of intellectual development in early childhood: Final Report. University of Illinois, Urbana, June, 1957.

Berger, Barbara. Researching Dotteressa Montessori. The American Montessori Society Bulletin, 1968, 6, No. 1.

Berger, Barbara. A longitudinal investigation of Montessori and traditional prekindergarten training with inner-city children: a comparative assessment of learning outcomes. Center for Urban Education, 1969.

Goodman, Mary E. Race Awareness in Young Children, Addison-Wesley, 1952, pp. 258-260.

Grotberg, Edith (ed). Critical Issues in Research Related to Disadvantaged Children, Educational Testing Service, 1969.

Hertzig, Margaret E., Birch, H. G., Thomas, A. and Mendez, Olga. Class and ethnic differences in the responsiveness of preschool children to cognitive demands. Monogr. Soc. Res. Child Developm. 1968, 33(1), Serial No. 117.

Jensen, Judith and Kohlberg, L. Report of a research and demonstration project for culturally disadvantaged children in the Ancona Montessori School. Submitted to Office of Economic Opportunity, August, 1966.

Karlson, A. L. A naturalistic method for identifying behavioral aspects of cognitive acquisition in young children participating in preschool programs. Ph.D. Dissertation, University of Chicago, in preparation, 1970.

Marshall, Helen and McCandless, B. A study in prediction of social behavior of preschool children. Child Developm., 1957, 28, pp. 149-159.

Stodolsky, Susan and Jensen, Judith. Final Report: Ancona Montessori research project for culturally disadvantaged children. Submitted to Office of Economic Opportunity, September, 1969.

Wechsler, D. Wechsler Preschool and Primary Scale of Intelligence Manual. The Psychological Corporation, 1967.

Zigler, E. and Butterfield, E. C. Motivational aspects of changes in IQ test performance of culturally deprived nursery school children. Child Developm., 1968, 39(1), 1-14.

APPENDIX A

Ancona School Headstart Program Ratings of Behavior During Individual Intelligence Testing

Name ____________________ Date __________ Stanford-Binet Form ______ Examiner ______________

Distractibility

1	2	3	4	5	6	7	8	9
Completely absorbed by task. Maintains interest throughout, remains oriented to E between items		Interested & attentive, with little attention to things external to the test		Normal attentiveness. Tasks elicit sufficient attention, though attention may occasionally wander between items		Attracted by things external to the test, but can return to task. If child tries to maintain attention, it is with some effort		Difficult to get and hold attention

Activity Level

1	2	3	4	5	6	7	8	9
Hyperactive; activity out-of-bounds		Quite active, but with control of own activity		Fairly active, but able to sit quietly for the testing		Rather sedentary, slow; or somewhat constricted in action (Circle which)		Depressed or very constricted (Circle which)

Speed of Response - Verbal Items

1	2	3	4	5	6	7	8	9
Responds very rapidly to instructions		No hesitation in responding		Some deliberation, but responses not generally slow		Usually slow to respond		Slow to respond; urging needed

Speed of Response - Performance Items

1	2	3	4	5	6	7	8	9
Responds very rapidly to instructions		No hesitation in responding		Some deliberation, but responses not generally slow		Usually slow to respond		Slow to respond; urging needed

Initiative in Dealing with Test Materials

1	2	3	4	5	6	7	8	9
Impulsively handles materials, begins own tasks		May begin to handle materials, may feel he knows what to do		Waits for instructions, but eager to begin		Responds to instructions, but does not initiate activity; is not "set" to begin until instructions given		Urging needed to respond

Self-Confidence on Tasks

1	2	3	4	5	6	7	8	9
Over-confident; does not recognize own limitations		Quite confident in own ability. Answers assertively		Realistically self-confident. May show recognition of own limitations; responds matter-of-factly		Some distrust of own ability, hesitancy		Distrusts own ability. Hesitant in response; may express concern about adequacy of responses

Persistence

1	2	3	4	5	6	7	8	9
Can't give up, even after much effort		Very persistent; difficult tasks lead to redoubled effort		Some persistence; doesn't give up without trying		Tendency to give up after first attempt unsuccessful		Gives up easily, when answer does not come almost immediately

Reaction to Failure

0	1	2	3	4	5	6	7
Ignores failure; success or failure not an issue	Aware of failure, but shows no discomfort, and does not need reassurance		Some discomfort at failure can be seen, but confidence easily restored		Disconcerted by failure, needs frequent reassurance		Withdrawing, hostile or denying (Circle which)

Sense of Intellectual Challenge

1	2	3	4	5	6	7	8	9
Hard tasks elicit greater interest and a sense of challenge		Hard tasks are met with special effort		Effort expended on problem is appropriate to level of difficulty, but there is no sense of challenge		Apparently more comfortable with easy tasks		Prefers only easy tasks

Willingness to Continue with Test

1	2	3	4	5	6	7	8	9
Eager to continue		Maintains active interest		Mild interest in tasks; continues to try		Loses interest in tasks, but continues to comply		Actively seeks termination

Fear of Adult

1	2	3	4	5	6	7	8	9
No shyness; quite self-assured		Rather confident		Neither confident nor fearful		Rather timid		Painfully shy, constricted

Social Initiative with Adult

1	2	3	4	5	6	7	8	9
Attempts to dominate the situation		Often initiates social interchange		Responsive, but usually does not initiate social interchange		Rather passive		Never takes initiative; responses minimal

Communication of Affect

1	2	3	4	5	6	7	8	9
Almost no inhibition of affective expression		Expresses affect freely, but with self-control		Occasional expression of affect		Tends to inhibit affect, or rather flat (Circle which)		Very flat; no emotional expression

Compliance with Adult

1	2	3	4	5	6	7	8	9
Extremely sensitive to adult's wishes; constantly looks to adult for permission to act		Tends to ignore own needs; rather anxious to comply		Makes own needs known, but quite willing to comply		Somewhat unwilling to comply		Actively negativistic

Verbalization

1	2	3	4	5	6	7	8	9
Speech perfectly clear and understandable		Occasional errors within generally good speech		Speech adequate; there are errors, but speech is still easily understood		Speech sometimes difficult to understand. This may be true especially when speaking rapidly		Speech very difficult to understand

Descriptive comments (a sentence or two, with examples if possible):

Articulation:

Structure:

Vocabulary:

APPENDIX B

Report on Social Work Program

by Lila Gordon

The social work program covered a wide range of services which were begun, as in previous years, on an attempted group basis. The assistant director of the school, a trained worker, set up weekly parent meetings which included active professional and lay community people from programs and agencies attempting to meet needs applicable to Ancona parents, e.g., the principal of the neighborhood public school to which most of the children will eventually transfer. However, the meetings were poorly attended despite individual contact with the families to encourage their participation. The one successful exception was an overflow group of both Ancona and non-Ancona neighborhood residents who came to hear a Department of Urban Renewal representative speak to the current efforts to help poor people purchase homes. However, follow-up to this meeting by the social worker with three individuals who were thus motivated resulted only in frustrating efforts by worker and clients to traverse bureaucratic and political obstacle courses. Although the Ancona group work program was of dubious value, it was clear that the parents were active still in groups of their own choosing: National Welfare Rights Organization, Operation Breadbasket, the Hyde Park Community Conference, the local park breakfast program.

The director of the school, also a trained social worker, had been administrative head of the school since 1967 and had had regular contact with most of the families as individual needs indicated. Crisis intervention continued to be paramount. It is clear that inadequate welfare assistance, non-existent lower rental livable housing, police-poverty population interaction, and inter-black conflict are fundamental to the entire reality functioning capacities of all the families. To cite a few specific daily coping concerns: no welfare budget for children's winter clothing was allocated in an inflation-battered grant; only ten units of a newly-constructed 250 apartment neighborhood project were allotted to low-income families ("and we will consider only those with good references"); parents were questioned by police to determine children's possible involvement in area killings, holdups, purse-snatching, fire-bombings, break-ins; gangs impinged increasingly on younger children. Especially heard was the increasing ambivalence about black militants and conflict as to the black faction with which they could or wanted to identify. A Black Parents' Group emerged from the middle-class parent body at the school during the year in an effort to seek possible funding in the black community for the dying Headstart program; the participation of the Headstart parents was solicited by this group with minimal response.

However inadequate or disappointing the social work concrete results appear, one fact stands out clearly: the Headstart children came to school regularly, and parents were strongly identified with Ancona as the school which their child or children attended and where they wished them to continue to attend. From social workers' observation, teachers continued to be very highly motivated with the children concerned about their learning, emotional, and psychological needs, contacting parents when children did not attend school, and offering concrete assistance (not financial) on a voluntary basis. The children were truly integrated into the classes with children they had known for several years.

The most concrete possible demonstration of parents' wish for children to attend Ancona is that for the first year, it was not necessary to pursue parents for re-registration for the following year: every child but one (who was about to move from the area) was re-registered for the 1970-71 school year, even though the entire parent body and staff was aware that Headstart funding would be terminating. In addition, ten unsolicited Headstart applications were received from new neighborhood parents. The board of directors of the school voted at its July, 1970, meeting to assume funding responsibility for Headstart children at the same dollar amount granted by OEO for tuition in 1969-70.

At this writing, thirty-nine Headstart children are registered for the 1970-71 school year, sixteen at the elementary level, twenty-three in the preschool. The tuition dollar amount, at increased school fees, is approximately $30,000 which the board of directors hopes to fund through its direct efforts. This major commitment reflects the determination of this parent-owned school to continue as a community-based institution including representatives of all socio-economic groups in the area. Philosophically, the school has rejected the idea of any continued funding through local Headstart and Follow-Through sources because of the requirement of having 80% disadvantaged population in one classroom: the children will remain fully integrated throughout the entire school.

With the completion of the research and demonstration program, formal testing, observation, and follow-up of the children will be ended. However, five years of Headstart programming at Ancona has firmly established a climate of interest, commitment, and earnest involvement in the program. Enrollment will continue as an Ancona-funded program (or deficit) in 1970-71. The director of the school has resigned to assume another position but will remain social work consultant for the school on a voluntary basis; contact with the families will be continued by the teachers with the social worker participating as need arises. Hopefully, some effort will be made by the staff to continue follow-up.

Communication from all teachers, office staff, and many middle-class parents indicates that however adequate a growth experience Headstart may have been for the children and families in the program, everyone else participating has testified to experiencing exciting professional and/or personal growth. The report of the last year of the program must emphasize, underline, and proclaim the repeated first hand confrontation with the uncompromising and unremitting evidence: all goals for which all compensatory educational programs have striven can be achieved only if families have a guaranteed cash income and services adequate to meet the legitimate human needs based on maximizing dignity and opportunity for all individuals in our society.